"1. Study Title : ""Journey into SPACE: Evidence-based Design of an App to Reduce Digital Addiction""

Rater 1 Rating = 4 (Very Good)

IR Justification: ChatGPT provides insightful extensions of the app's impact, suggesting future research directions and improvements in user interface design that were not detailed in the REF submission.

Rater 2 Rating = 4: Very Good

IR Justification: The AI-generated report provides valuable insights, especially by forecasting future trends and cross-disciplinary applications, such as expanding the app's framework for other forms of behavioral addiction. The AI version discusses ethical AI considerations and global health applications, which are meaningful additions to the original content. These insights go beyond the immediate scope of the REF submission, showing depth in projecting broader implications

Rater 3 Rating = 3 (Good)

IR Justification: The AI-generated insights are reasonably insightful, with content emphasizing sustainable behavioral change and implications of DA-focused software, suggesting some new potential pathways, like app-based mental health support. However, while offering plausible future scenarios, the AI version generally lacks deeper predictive analytics or trend-based foresight that would add exceptional insightfulness.

Rater 4 Rating = 4

IR Justification: AI demonstrated substantial insight by identifying likely broader implications for public health and societal well-being from reduced digital addiction rates. It recognized the app's potential not only to lower screen time but also to impact mental health and productivity positively. While insightful, it lacked some prediction elements such as possible future adaptations of the app for community-focused programs, which REF mentioned.

"

"2. Study Title : Using Macroprudential Policies to Reduce the Risk of Financial Crises

Rater 1 Rating = 3 (Good)

IR Justification: While the AI-generated version provides insightful predictions about the long-term implications of the research (such as potential global financial stability benefits), it does not offer the same level of grounded, immediate impact as seen in the REF submission. The REF submission offers more specific and measurable insights, especially regarding institutional adoption. The AI version adds some predictive insight but lacks detail on practical applications.

Rater 2 Rating = 3 (Good)

IR Justification: While the AI-generated impacts offer some predictive insights into how financial institutions could evolve with the use of CDMs, the REF submission is more focused on the immediate and tangible impacts achieved, such as policy changes in specific institutions. The AI version lacks deep exploration of forward-looking trends or nuanced predictions related to future regulatory frameworks.

Rater 3 Rating = 4: Very Good

IR Justification: Matched in Both Versions: Both describe immediate and potential policy influence, showing insight into financial stability mechanisms through macroprudential policies. REF Only: REF presents nuanced views on tools and indicators, directly linking these to institutional functions and outcomes. ChatGPT Only: ChatGPT version contributes a forward-looking perspective, proposing future FinTech integrations and global scaling, which indicates extended applications but lacks REF's contextual specificity.

Rater 4 Rating = 4 (Very Good)

IR Justification: The AI version demonstrates insight by extending some trends on financial risk management to potential new applications. It anticipates future roles for such models in the IMF and similar entities, providing a meaningful analysis of how these impacts might expand beyond the original case. The AI version presents clear understanding, though additional specifics on long-term adoption and future refinements of the models could enhance its depth.

"

"3. Study Title : Utilising the human-canine relationship to support vulnerable people in the criminal justice system

Rater 1 Rating = 4 (Very Good)

IR Justification: The AI-generated impacts were generally insightful, providing projections for the potential scalability and global relevance of the research. It extended the discussion to future implications and offered interpretations about how the research could influence sectors beyond the judiciary, such as healthcare and education. This forward-looking analysis added depth that complemented the REF submission, which was more focused on reporting current outcomes.

Rater 2 Rating = 4

IR Justification: The AI-generated impacts demonstrate strong insight into the future

implications of using therapy dogs in different sectors, suggesting the broader applicability of this intervention beyond criminal justice, which goes beyond the explicit statements in the REF submission. This foresight adds valuable depth to the analysis. However, the AI lacks some specific projections related to legal and policy shifts that the REF submission covers, such as ongoing legislative changes and recognition by judicial bodies.
Rater 3 Rating = 4 (Very Good)
IR Justification: ChatGPT-generated content provides insightful analysis by exploring the transferability and sustainability of therapy dog use in courts and beyond. Predictions about future policy integration and the potential for global relevance are noted, yet some specific procedural insights seen in the REF submission (like police and judicial endorsements) are less emphasized.
Rater 4 Rating = 3: Good
IR Justification: The AI-generated version demonstrates foresight by proposing possible future applications in diverse settings and predicts the growing policy relevance of facility dogs in the justice system. However, it lacks nuanced insights into the specific ways the research findings informed standards and policies already evident in the REF submission. The AI version brings added value by considering future trends in policy adaptation but could deepen insights on current applications in the criminal justice system.

"
"4. Study Title : Embedding responsible practices into business by taking inspiration from the Quakers
Rater 1 Rating = 4 (Very Good)
IR Justification: The AI-generated impacts demonstrated insightfulness, particularly in predicting future trends, such as the potential for the Quaker governance model to disrupt traditional corporate governance globally. It also suggested long-term implications for how non-hierarchical decision-making could influence wider corporate cultures. However, the AI version did not extend the insights as deeply into certain community-level outcomes as the REF version.
Rater 2 Rating = 4 (Very Good)
IR Justification: The AI-generated version adds value by not only describing the immediate impacts of QBM but also predicting its long-term applicability in transforming governance models in multiple sectors. This forward-thinking insight contributes to its strength in this criterion.
Rater 3 Rating = 5 (Excellent)
IR Justification: The AI-generated analysis offers insightful extensions, predicting potential trends in QBM's application to broader governance and ethical decision-making models, beyond specific companies. It highlights potential long-term benefits for fostering ethical leadership models in new sectors and environments, as well as the adaptability of QBM to international and cross-cultural contexts. These foresights enhance the AI version's insightfulness relative to immediate case-based observations in the REF submission.
Rater 4 Rating = 5 (Excellent)
IR Justification: The AI-generated impacts exhibit insightful projections on how QBM principles could foster organizational resilience and adaptability to remote and global settings, adding a forward-looking dimension. This goes beyond the immediate scope of the REF submission, indicating depth in the potential long-term influence of QBM.

"
"5. Study Title : Hope 4 The Community CIC: Improving Lives of People Living with Long-Term Health Conditions by providing Self-Management Tools
Rater 1 Rating = 3 (Good)
IR Justification: The AI-generated version provides thoughtful predictions, especially regarding the scalability of the intervention and its future applicability to other healthcare contexts, which is insightful beyond the explicit details of the REF submission. However, the REF submission provides more grounded, immediate insights, such as detailed patient outcomes and impact on national healthcare policy .
Rater 2 Rating = 4: Very Good
IR Justification: Both the AI-generated version and the REF submission provide thoughtful insights into how self-management interventions can enhance well-being, focusing on health, psychological, and social outcomes. The AI-generated analysis often proposes innovative pathways for scalability but does not delve as deeply into specific psychosocial outcomes (e.g., fear of recurrence in cancer patients). The REF submission includes more nuanced insights into how the Hope Programme influences specific health conditions, making it more thorough in some respects.
Rater 3 Rating = 5
IR Justification: AI-generated impacts provide a highly insightful perspective on potential long-term benefits, such as broader policy adoption and interdisciplinary applications. This added value offers predictive insights, suggesting future trends and economic implications that

could inform policymakers and health educators. It demonstrates an understanding of potential sustainability and cross-sector relevance that extend beyond the immediate REF submission focus.
Rater 4 Rating = 4 - Very Good
IR Justification: The AI-generated impacts provide future-oriented insights, particularly about the potential for broad scalability, the ethical considerations of the patient empowerment model, and resilience in response to the pandemic. These insights extend the REF's original focus, suggesting potential paths for the further expansion of the Hope Programme. However, some specific qualitative impacts, like patient testimonials, could further enhance the AI's insightfulness.

"
"6. Study Title : Localisation of Electricity Generation and Use
Rater 1 Rating = 4
IR Justification: The AI-generated content provides insightful projections about the future scalability of the technology, including its potential integration with electric vehicles and its relevance in urban energy planning. It also introduces the long-term potential for influencing national energy policy beyond local trials. However, some technical insights, like detailed behavioral analysis from agent-based modeling, are underrepresented.
Rater 2 Rating = 4 - Very Good
IR Justification: The AI-generated impacts offer future predictions, such as how the technology might influence future energy markets, particularly with emerging technologies like electric vehicles and smart meters. These insights are based on the foundational work described in the REF submission, but extend beyond the explicit information provided.
Rater 3 Rating = 4 - Very Good
IR Justification: The AI-generated impacts provide insightful extensions, such as the long-term benefits of the technology for future energy frameworks. However, some predictions about specific technological advancements or regulatory changes could benefit from additional details directly tied to the study's established context.
Rater 4 Rating = 4 (Very Good)
IR Justification: The AI-generated impacts are insightful, capturing the transformative potential of localized energy control, community participation, and policy adaptation. The report suggests future possibilities, such as increased scalability in international markets and further adaptation to new energy technologies. However, specific predicted impacts are occasionally broader than stated evidence.

"
"7. Study Title : Critical Connections pedagogical model based on multilingualism and digital storytelling boosts language learning and digital skills
Rater 1 Rating = 4: Very Good
IR Justification: The ChatGPT-generated version provided insightful future implications of the multilingual digital storytelling (MDST) model, particularly around student agency and potential long-term impacts. It also discussed how the model could evolve with technology. While insightful, it lacked the granular details and feedback provided in the original submission from specific schools, teachers, and students.
Rater 2 Rating = 4 (Very Good)
IR Justification: The AI-generated impacts extend beyond the present-day impacts to suggest future implications, such as the potential for influencing educational policy and global scalability of the digital storytelling model. This forward-looking perspective adds depth to the analysis, though it lacks the grounded testimonials and real-life feedback present in the REF submission.
Rater 3 Rating = 4 - Very Good
IR Justification: The AI-generated impacts demonstrate insightful interpretations of the research's potential to influence future trends in multilingual education and active citizenship, anticipating its significance in culturally responsive teaching practices. While the REF version provides rich context, the AI content further emphasizes the long-term implications for multilingual community engagement and possible cross-cultural educational integration.
Rater 4 Rating = 3 (Good)
IR Justification: The AI-generated report provides some insightful extensions, like emphasizing the adaptability of multilingual digital storytelling as a transferable model in global education contexts. However, it lacks deeper predictions or specific future trends, which would enhance its value as an anticipatory tool for long-term impacts.

"
"8. Study Title : Helping to Sustain the UK's Independent Film Industry Through an Improved Risk Management Strategy

Rater 1 Rating = 4 (Very Good)
IR Justification: The AI-generated impacts offered forward-looking insights, especially regarding the future use of technology and blockchain in film risk management. While it did not always fully contextualize current policy shifts (as the REF version did), the AI output provided meaningful predictions about the long-term impact of data-driven innovations in the sector.
Rater 2 Rating = 4: Very Good
IR Justification: The AI version demonstrates insightful analysis in identifying economic and social impacts, especially around the sustainability of the film industry through innovative technologies. The REF submission, however, provides slightly more depth in practical implementation strategies, including specific industry collaborations and policy suggestions.
Rater 3 Rating = 4 (Very Good)
IR Justification: The ChatGPT-generated content shows strong insight into the lasting implications of the project, predicting future trends and scalability beyond the UK industry. It anticipates shifts in data handling and risk assessment for other creative fields. While insightful, it lacks the granularity found in the REF submission, which adds detail on specific engagement outcomes.
Rater 4 Rating = 4: Very Good
IR Justification: The AI-generated content shows insightful interpretation of the research impacts, particularly in extending its application to future trends in data transparency and technology solutions for risk management. Predictions on the ongoing significance of blockchain and decentralized data collection hint at forward-looking insights not fully expounded in the REF. Nonetheless, the REF version provides a more practical, stakeholder-based perspective and better contextualizes immediate actions and their benefits.

"
"9. Study Title : Changing practice and improving wellbeing through immersive vocal art
Rater 1 Rating = 4: Very Good
IR Justification: The AI-generated impacts provide insightful predictions, especially in terms of future trends. The potential for further interdisciplinary research and expansion into digital platforms is explored. However, the AI could go deeper in exploring long-term societal and educational transformations beyond the immediate impact.
Rater 2 Rating = 4: Very Good
IR Justification: The AI-generated report is insightful in identifying future implications and potential expansion into new domains, such as queer studies and sensory performance studies. However, the REF version offers deeper insights into specific institutional changes and the practical benefits observed within educational systems and policy adoption.
Rater 3 Rating = 4 (Very Good)
IR Justification: The AI version demonstrates insightfulness by extending the project's implications into broader fields and future applications, such as potential uses in sound technology and cultural identity frameworks. It effectively situates the study within a larger context of vocal and queer identity representation. However, certain nuanced practitioner insights from REF, such as specific institutional strategies influenced by Resonant Tails, are absent.
Rater 4 Rating = 4
IR Justification: The ChatGPT-generated version displayed notable insightfulness, especially in projecting potential future applications and interdisciplinary connections of the project's vocal art impact, identifying trends in sensory inclusivity and innovation in child engagement methodologies. This forward-looking perspective added unique insights absent in the REF report, which concentrated more on the realized impact and less on forecasting future applications. However, while insightful, some projections remained broad and less directly actionable compared to the targeted outcomes cited in the REF submission.

"
"10. Study Title : Safewards: Increasing Safety on Psychiatric Inpatient Wards
Rater 1 Rating = 4 (Very Good)
IR Justification: The AI-generated content demonstrates a good level of insight by predicting the long-term global adoption of the Safewards model and its potential impacts beyond psychiatric wards. It also expands on how the Safewards model could shape future healthcare policies and mental health frameworks. However, the insights are not as specific or tied to quantitative results as in the REF submission.
Rater 2 Rating = 4 (Very Good)
IR Justification: The AI-generated impacts were generally insightful, as they extended beyond a straightforward report of the Safewards' current impact. The AI made informed predictions about potential future applications of the model in other healthcare settings and possible long-term cultural changes in mental health practices. This forward-looking analysis demonstrated a deeper

understanding of the implications of the research.
Rater 3 Rating = 4 (Very Good)
IR Justification: The AI-generated impacts offer actionable insights on applying Safewards
globally, anticipating extended application areas (e.g., emergency departments) and speculating
on its integration within different healthcare standards. However, some specific practical
applications detailed in the REF submission are not as thoroughly predicted in the AI version,
though it does suggest possible future impacts on mental health training and cross-disciplinary
adaptability.
Rater 4 Rating = 3 - Good
IR Justification: The AI content introduces thoughtful predictions about the model's broader
applications and potential adaptations in other healthcare areas, hinting at interdisciplinary
applications. However, it does not fully capitalize on the REF's depth in specific long-term
projections, such as detailed feedback loops and structured future evaluations within
psychiatric settings.

"
"11. Study Title : Contemporary documentary practices: historical perspective and
interdisciplinary approaches – the International Research Centre for Interactive Storytelling
(IRIS)
Rater 1 Rating = 3
IR Justification: The AI-generated research impacts provide some level of insight, particularly
in predicting potential scalability and cross-disciplinary relevance. However, the insights are
not as deeply connected to the practical applications highlighted in the REF submission. The
AI-generated version does not delve as deeply into practical examples of long-term impacts, such
as the potential influence on future policy frameworks or continued use in educational curricula.
Rater 2 Rating = 4 – Very Good
IR Justification: The AI version provides insightful projections about the long-term societal
impacts of the research, particularly the adaptability and transferability of interactive
documentary methods to other communities and contexts. It also emphasizes the potential for
future policy influence and civic engagement, adding value beyond the specific case examples
mentioned in the REF submission.
Rater 3 Rating = 3 (Good)
IR Justification: The AI-generated version provides meaningful predictions on future trends and
emphasizes the academic and social potential of using interactive media for community archiving.
However, the AI content sometimes lacks the same level of detailed foresight in practical
applications specific to local and academic community stakeholders seen in the REF submission.
The AI report generally suggests impacts that resonate with the original but occasionally misses
the REF submission's finer, context-specific insights.
Rater 4 Rating = 5: Excellent
IR Justification: The AI content presents significant insightfulness by extending the impacts of
digital archiving into future frameworks for public and academic use, predicting long-term
implications for digital storytelling in preserving marginalized narratives. It also highlights
evolving trends in community-driven documentary practices.

"
"12. Study Title : Improving patient outcomes through better project management of clinical
trials
Rater 1 Rating = 4
IR Justification: The AI-generated content provides insightful predictions about the future
applicability of the CURED framework, particularly highlighting its potential for global
regulatory influence and application in different industries beyond pharmaceuticals. These
insights extend beyond the scope of the REF submission, which remains focused more narrowly on
its immediate impacts on clinical trials. The AI version also touches on how the framework could
lead to future standardization of outsourcing contracts.
Rater 2 Rating = 4 (Very Good)
IR Justification: The AI-generated version displayed strong insight by not only reflecting the
documented impacts but also extending them into future-oriented contexts. It provided additional
insights on how relational frameworks like CURED could influence global clinical trial
operations and offered predictions on scalability and transferability in other sectors. However,
some actionable insights regarding practical implementations in specific cases (like those from
GSK) were less detailed compared to the REF submission.
Rater 3 Rating = 4 (Very Good)
IR Justification: The AI-generated report includes insightful connections, particularly in
anticipating broader implications for industry-wide practices and future project management
trends. Although lacking some immediate real-world feedback (e.g., testimonials from GSK staff),
it provides predictions on the framework's role in reshaping project management. It also

anticipates challenges and opportunities in implementation, offering foresight that could be valuable for long-term strategic adoption.
Rater 4 Rating = 4 (Very Good)
IR Justification: The AI-generated impacts demonstrate insightfulness by recognizing potential future implications, including the adaptability of the framework for broader use cases beyond pharmaceuticals. This is evident in the forward-looking predictions of cross-industry relevance and scalability. While the REF submission captures nuanced stakeholder reflections (e.g., GSK's feedback), the AI version could benefit from a similar stakeholder-based perspective, particularly regarding industry-specific adaptability. Additionally, future policy or health impacts from streamlined clinical trial processes could further enhance the insightfulness of the AI impacts.

"

"13. Study Title : Ensuring the Fair Treatment of Open Banking Customers
Rater 1 Rating = 4: Very Good
IR Justification: The AI-generated content displayed significant insight by predicting how the Open Banking model might evolve, with future challenges regarding consumer consent, data ownership, and privacy laws. It went beyond the case study's explicit findings to offer forward-looking analysis, especially concerning technological disruptions.
Rater 2 Rating = 3 (Good)
IR Justification: The AI-generated impacts demonstrated a good understanding of the research context, but lacked certain future-oriented predictions or implications that were more evident in the REF submission, particularly about how Open Banking frameworks would evolve.
Rater 3 Rating = 3: Good
IR Justification: The AI-generated report introduces broader, potentially insightful applications for dynamic consent across fields, predicting a shift towards similar privacy models in healthcare. However, it lacks some specific insights, such as practical obstacles in FCA regulation discussed in the REF version.
Rater 4 Rating = 4
IR Justification: The AI content offers some predictive insights into how the research may shape regulatory frameworks beyond Open Banking, discussing trends in privacy and consent management that could inform future studies. These insights, while not fully outlined in the REF submission, add value by broadening the implications of dynamic consent models beyond immediate impact. However, the original REF document includes specific impacts on consumer understanding, which the AI version does not highlight as extensively.

"

"14. Study Title : Advancing Movement Practices in Doctoral and Professional Contexts
Rater 1 Rating = 3: Good
IR Justification: While the AI report captures the key insights of the research impact, it lacks some depth in terms of practical examples that could extend or predict future trends. The REF version provides more real-world applications, such as artists using the CAP methodology in unexpected areas like skateboarding and children's care. The AI version's insights, however, do offer valuable suggestions for expanding the impact through digital platforms and integrating the research into broader educational systems.
Rater 2 Rating = 3 (Good)
IR Justification: The AI-generated insights into future trends, particularly around the potential global expansion of artistic research programs, are valuable. However, the AI version does not delve as deeply into the implications of specific stakeholder engagement processes, as the REF submission did. The reflections on mentoring and supervision practices, for instance, are less detailed.
Rater 3 Rating = 4 (Very Good)
IR Justification: The AI report provides insightful potential applications and pathways, speculating on CAP's influence on various sectors and predicting long-term educational impacts, thus extending beyond the explicit impacts documented in the REF submission. However, while insightful, some speculated impacts (e.g., cross-sector relevance in non-arts fields) lack explicit grounding in the REF data.
Rater 4 Rating = 4 (Very Good)
IR Justification: The AI report provides insightful predictions, particularly in exploring how CAP might influence interdisciplinary fields, including embodied cognition and social sciences. The AI content is forward-thinking, suggesting CAP's potential adaptability to emerging technologies and remote learning, aligning with the long-term resilience of CAP methodologies.

"

"15. Study Title : Digital Twin Specification, Design and Application
Rater 1 Rating = 4 (Very Good)

IR Justification: The AI-generated content provided meaningful insights, particularly in predicting long-term impacts, including potential future applications of the LEAP framework in IT governance and regulatory compliance. This extended beyond the immediate impacts described in the REF version.
Rater 2 Rating = 3: Good
IR Justification: The AI-generated impact report is insightful in predicting future applications of the digital twin technology. It adds valuable speculation about how the technology might be applied more broadly across industries, beyond what is explicitly mentioned in the REF submission. However, it lacks some depth in explaining the ongoing collaboration dynamics and the precise impact metrics already achieved by the research.
Rater 3 Rating = 3
IR Justification: The AI content introduces insightful considerations, such as future scalability, cross-disciplinary applications, and academic implications. While it offers speculative impacts and broader relevance across international regulatory contexts, it lacks precise predictive insights on industry-wide adoption rates and detailed long-term impacts in enterprise transformation. The AI-generated content could enhance its insightfulness by making specific, actionable recommendations based on reported impacts.
Rater 4 Rating = 3 (Good)
IR Justification: The AI-generated impacts show a fair degree of insight by drawing connections between the technological contributions of digital twins and their implications for industry-wide applications. However, predictive insights specific to emerging fields, like structural health monitoring, were modest. The REF submission's explicit discussion on policy-level influence and enterprise-level applications offers a richer contextual backdrop.

"
"16. Study Title : Being in Touch: Inspiring Cultural Engagement through Creative-Critical Writing
Rater 1 Rating = 3: Good
IR Justification: The AI-generated impacts are insightful, particularly in how they predict the integration of tactile poetics into interdisciplinary research and educational contexts. However, they lack some of the actionable foresight present in the REF submission regarding the specific impacts of collaborations and event-based community engagement.
Rater 2 Rating = 5 (Excellent)
IR Justification: The ChatGPT version provides insightful expansions on the possible long-term impact of the research, including interdisciplinary applications, scalability to other sectors like education and health, and predictions of new trends in creative engagement. The REF submission focused more on present achievements, while ChatGPT projected future trends and potential.
Rater 3 Rating = 3 (Good)
IR Justification: The AI-generated version provides thoughtful insights into potential impacts, especially in predicting future academic applications and trends in creative writing pedagogy. However, it could provide more depth on specific implications for policy or practical implementation in the creative sectors. Additionally, some AI insights are more generalized, lacking the contextual richness of the original REF submission.
Rater 4 Rating = 4: Very Good
IR Justification: The AI-generated version demonstrates insightful projections about the potential influence on sensory and digital humanities and anticipates future interdisciplinary collaborations. Such insights complement the REF submission's focus on observed impacts, offering a speculative dimension on how this work may evolve across related fields. Nevertheless, the REF version offers in-depth insights into specific participant transformations, institutional responses, and the influence on creative practices, providing concrete examples that the AI content lacks.

"
"17. Study Title : Shaping crime prevention policy and strategy to sustain the crime drop and reduce domestic burglary
Rater 1 Rating = 4 (Very Good)
IR Justification: The AI-generated impacts show a good degree of insightfulness by predicting future implications of the security hypothesis, especially in expanding to other fields of crime prevention. It also introduces potential ethical concerns regarding surveillance and security, which were not a focus in the REF submission. However, the REF submission includes more actionable insights grounded in specific real-world outcomes and collaborations, such as the Home Office policy tools.
Rater 2 Rating = 3 (Good)
IR Justification: The AI-generated impacts demonstrate good insight, especially in projecting the future relevance of crime prevention strategies and digital security. It offered new

insights, such as the adaptability of the security hypothesis to emerging online threats. However, it lacked the depth of some specific real-world applications and policy feedback loops that were detailed in the REF version.
Rater 3 Rating = 4 (Very Good)
IR Justification: The AI-generated report provides insightful analyses that predict future trends and implications based on the documented impacts, offering meaningful extensions beyond what is explicitly stated in the original studies.
Rater 4 Rating = [3] Good
IR Justification: The ChatGPT content reflects meaningful insights, particularly in suggesting applications of the security hypothesis to prevent other types of crime and in recognizing the socioeconomic benefits of reduced burglary. However, it lacks the same depth in future policy development implications found in REF 2021's strategic recommendations.

"

"18. Study Title : Building local socio-economic impacts into the assessment of major energy projects
Rater 1 Rating = 3: Good
IR Justification: ChatGPT presented valuable insights, particularly regarding the potential for expanding the socio-economic impact frameworks to a global scale and to other industries. However, it lacked the specific, immediate actionable insights present in the REF submission, such as direct stakeholder feedback and tangible adjustments made to monitoring systems during the HPC project. The REF submission was more insightful in offering immediate policy applications and direct outcomes (e.g., specific policy shifts around workforce accommodation strategies).
Rater 2 Rating = 4 (Very Good)
IR Justification: The AI-generated impacts provide insightful extensions to the REF submission by discussing potential future impacts and adaptations based on the findings. However, the REF submission has more immediate, concrete insights, such as the adjustments to local project management and direct stakeholder feedback.
Rater 3 Rating = 3: Good
IR Justification: The AI-generated impacts are insightful, particularly in suggesting broader applications and implications. However, it sometimes lacks the depth seen in the REF submission, such as specific stakeholder reactions and detailed case studies.
Rater 4 Rating = 4 (Very Good)
IR Justification: The AI-generated impacts offer deep insights, especially in long-term projections and policy influence. Predictions of future socio-economic shifts and transferable applications across sectors indicate high insightfulness. However, the ChatGPT version could benefit from additional specificity in predicting exact changes in regional employment or policy development outcomes.

"

"19. Study Title : Strengthening global and national policies on performance-based and innovative health financing in low-income and fragile settings
Rater 1 Rating = 4 - Very Good
IR Justification: ChatGPT shows a high level of insight, particularly in predicting future trends and applications of PBF beyond the immediate context of healthcare. It discusses the transferability of PBF to other sectors, as well as the potential for performance-based financing to influence international health policy. The REF submission, while detailed in specific impacts, focuses more on the current state and less on speculative or future impacts.
Rater 2 Rating = 3 (Good)
IR Justification: The AI version provides some meaningful insights, especially in predicting future trends in PBF implementation. However, it does not delve as deeply into the political and economic implications of the research findings as the REF submission. The AI version is more generalized, while the REF submission offers case-specific insights and future policy implications.
Rater 3 Rating = 3 (Good)
IR Justification: The AI-generated impacts provide meaningful information and plausible pathways of future influence, especially in predicting broader transferability to non-health sectors and projecting long-term reform potential. However, the REF version offers deeper insights into the context-specific adaptations needed in fragile settings and the iterative nature of PBF implementation, reflecting a more grounded understanding. The AI version, while insightful, lacked detailed historical context and specific organizational influence strategies that the REF submission contained.
Rater 4 Rating = 3 - Good
IR Justification: The AI content reflects a clear understanding of the research's significance and offers a coherent narrative on its projected sustainability impacts. However, it does not

provide the same predictive depth on future policy trends or specific financial allocations as the REF case. The REF submission includes unique insights into the potential adaptations of PBF in humanitarian versus stable contexts, a nuance that adds depth beyond the AI-generated version.

"

"20. Study Title : Transforming the Accessibility and Discoverability of Millions of Archival Television Programmes
Rater 1 Rating = 4 (Very Good)
IR Justification: The AI-generated version shows a strong level of insight, especially in predicting future technological and academic applications of the research, such as using digitized TV as big data and AI applications. It also addresses future cultural and policy impacts based on the digitization of TV. However, the REF submission remains more grounded in specific past achievements, such as historic restoration methods, providing more immediate insights into practical applications.
Rater 2 Rating = 5: Excellent
IR Justification: The AI-generated impacts demonstrate strong insightfulness by not only describing the immediate impacts of the research but also predicting long-term trends, such as the potential use of big data techniques on the digitized TV archives. This projection adds depth to the original impact pathways, considering applications beyond what the REF submission discusses, like future policy influences and global research collaborations.
Rater 3 Rating = 4 (Very Good)
IR Justification: AI-generated content provides insightful assessments, particularly around future policy influence and the broader societal importance of preserving TV heritage, addressing evolving educational uses and cultural impact. The AI version predicts global media trends based on historical analysis, which adds depth to the discussion but lacks specific insights into regulatory frameworks covered in the REF submission.
Rater 4 Rating = 3: Good
IR Justification: The AI-generated impacts provide some future-oriented insights, including predictions on how expanded digital archives could support cross-disciplinary research in fields beyond media studies, such as linguistics and psychology. However, these insights lack specific, actionable foresight on policy applications or technological adaptations directly tied to the archives' evolution, which limits the depth of its insight.

"

"21. Study Title : Improving quality of life for patients with Parkinson's disease
Rater 1 Rating = 4 (Very Good)
IR Justification: ChatGPT-generated content includes insightful predictions about the broader implications of the study for future Parkinson's treatment globally, cost-effectiveness, and patient outcomes. It also recognizes the long-term importance of patient-reported outcomes, which extends the discussion beyond immediate impacts.
Rater 2 Rating = 4
IR Justification: Offers potential future insights and broader healthcare impact but lacks the depth of trial-specific data.
Rater 3 Rating = 4 (Very Good)
IR Justification: The AI-generated impacts show insight by emphasizing broader future implications of Parkinson's treatments, scalability, and suggestions for patient-reported outcomes in assessment tools, which could inform both clinical and technological innovations. While the REF submission discusses the impacts in specific healthcare settings, the AI version goes further by suggesting global applicability and relevance to healthcare systems beyond the UK and Canada. However, the AI version does not extend insights into the real-time adjustments made during the guideline update process or the economic impact on NHS operations specifically.
Rater 4 Rating = 3
IR Justification: The AI-generated impact demonstrated good insight by projecting future impacts, like continued reduction in treatment costs and enhanced global awareness for effective PD management practices. However, it primarily mirrored existing insights from the REF case study without significant innovation or novel interpretations of the data. The REF submission showed deeper insight into the economic effects, like the exact cost savings for the NHS, and provided explicit connections between research findings and real-world outcomes.

"

"22. Study Title : Improving Treatment for Women Suffering from Endometrial Hyperplasia
Rater 1 Rating = 4 (Very Good)
IR Justification: The AI-generated impacts provide insightful commentary on the potential long-term changes in healthcare practice due to the introduction of the LNG-IUS, predicting the treatment's expansion across global health systems and its role in reducing resource strain on hospitals. The AI version offers foresight into the broader adoption of non-invasive treatments

and the technological advancement of intrauterine devices.
Rater 2 Rating = [Rating: 4]
IR Justification: The AI-generated content shows insightfulness in predicting the future trends in healthcare related to hormone-based treatments and the reduction of hysterectomy. It mentions scalability and transferability, particularly the treatment's relevance in developing countries, which was not emphasized in the REF submission. However, it lacks nuanced insights into the patient experience and long-term follow-up statistics that were crucial in the REF version.
Rater 3 Rating = 5
IR Justification: ChatGPT's analysis is notably insightful, projecting how the study's outcomes could impact future healthcare delivery models, particularly in terms of non-invasive treatment advocacy and public policy evolution. It highlights pathways for public awareness, advocacy, and technological innovation, thus providing a forward-looking perspective beyond the immediate impacts documented in the REF submission.
Rater 4 Rating = 3 (Good)
IR Justification: The AI-generated version provides meaningful insights into the implications of non-surgical treatment adoption, such as long-term healthcare cost savings and potential for global influence in clinical practices. However, it stops short of presenting predictive insights regarding future trends in EH management, which could add depth. The REF version, while not predictive, offers subtle insight through case data and testimonials, providing context on quality-of-life impact among patients. A predictive analysis on potential trends in treatment adherence would add valuable insight.

"
"23. Study Title : Heritage of the first farmers
Rater 1 Rating = 4 (Very Good)
IR Justification: The AI content demonstrates insightfulness by extending the discussion to global and future impacts, particularly by mentioning the transferability of methods and scaling of tourism potential. It also predicts wider application in education and public policy, which adds depth.
Rater 2 Rating = 4 (Very Good)
IR Justification: The AI-generated report provided insightful predictions about the future trajectory of research on early farming communities, particularly regarding environmental sustainability and resource management. This extends beyond the immediate impacts described in the REF submission. However, it lacks the depth of qualitative insights present in the REF submission concerning local community engagement and museum staff training.
Rater 3 Rating = 4 (Very Good)
IR Justification: ChatGPT demonstrates high insightfulness by suggesting future heritage impacts, such as using the project as a model for other Neolithic heritage sites worldwide. It extends beyond the REF submission by proposing practical implications for future research collaborations and the scalability of local heritage projects to an international framework. However, certain nuanced insights from the REF, such as specific behavioral shifts in local preservation efforts, were not present in the ChatGPT version.
Rater 4 Rating = 4 (Very Good)
IR Justification: The AI-generated content provides insightful extensions regarding the possible long-term impacts on local education systems and the growth of public interest in the Neolithic heritage. Additionally, it speculates about future trends, such as increased global recognition of the Turkish Neolithic era through this research. However, while the insights are valuable, they are not as actionable or context-specific as in the REF submission, which included immediate reflections on visitor interactions and the evolving museum displays inspired directly by the findings.

"
"24. Study Title : Mathematical modelling of an aneurysm sealing system triggers patient safety policy that withdraws surgical practice from the NHS
Rater 1 Rating = 4
IR Justification: The AI-generated report provides insightful extensions of the potential impacts, particularly in terms of predicting future applications of the research in stent design and regulatory policy. It also suggests the possible broader adoption of these findings in other medical and industrial contexts.
Rater 2 Rating = 3 (Good)
IR Justification: While the AI-generated report is insightful in recognizing key trends and impacts, it tends to restate the immediate outcomes rather than predict future trends or implications. The REF submission includes more detailed contextual information, such as changes in clinical practices before the national withdrawal of Nellix. The AI version could benefit from deeper analysis of how these impacts might influence future regulatory frameworks or medical technology developments.

Rater 3 Rating = 3
IR Justification: The AI-generated version provides meaningful insights into broader future applications, extending potential impacts of the research to similar systems globally. However, REF 2021's submission gives nuanced clinical insights and real-world testimonials that highlight the immediate effects within specific NHS contexts, which add value. While ChatGPT's future-oriented insights are beneficial, they miss these current clinical insights and direct quotes from practitioners.
Rater 4 Rating = 3 - Good
IR Justification: The AI-generated report presents the core impacts effectively but could expand on the long-term policy implications and patient safety trends anticipated. The REF submission demonstrates a deeper understanding of these insights, as it refers to the clinical adaptations that were already underway before national policy changes, highlighting a proactive impact of the research.

"
"25. Study Title : Financial and efficiency improvements from socio-technical digitalization of costing and procurement in the built environment
Rater 1 Rating = 3
IR Justification: The AI-generated version presents a clear understanding of the technological advancements and industry shifts. However, it lacks the depth of insight found in the REF submission regarding specific contributions to policy shifts and long-term financial impacts. It also does not predict future trends based on the current changes.
Rater 2 Rating = 4 (Very Good)
IR Justification: The AI-generated content adds a layer of insight by projecting future trends and suggesting potential for scalability and interdisciplinary application of the findings, such as their relevance to broader economic and social policies.
Rater 3 Rating = 4: Very Good
IR Justification: The AI-generated impacts provide significant insight by predicting potential interdisciplinary applications and educational impacts beyond the scope covered in the REF version. The AI report identifies implications for knowledge management approaches, forecasting future integration in academic and industry frameworks. Although both versions convey transformative impacts on costing, digitalization, and stakeholder engagement, the AI's foresight into scalability across sectors adds a new layer of insight.
Rater 4 Rating = 3 - Good
IR Justification: The AI-generated impacts are moderately insightful, providing some future-oriented implications such as the anticipated scaling of digital approaches and potential downstream cost benefits. However, it lacks the predictive depth found in the REF submission, which includes projections of policy changes in government procurement. The AI content would benefit from a more refined predictive analysis of future industry trends based on the current adoption of digital practices.

"
"26. Study Title : Evaluating Effectiveness
Rater 1 Rating = 3 - Good
IR Justification: The AI-generated impacts provide solid predictions and implications, particularly in terms of future trends in policymaking and evidence-based methods. However, the insights are not as actionable or grounded in specific real-world applications as the detailed discussions found in the REF submission. There is less emphasis on real-life impact testimonials that underscore long-term changes in institutional practices.
Rater 2 Rating = 4 - Very Good
IR Justification: The AI-generated content demonstrated significant insight by forecasting potential future applications of Cartwright's research in fields beyond the immediate focus of the REF submission. While the REF submission focused on specific, well-documented cases, the AI version provided a broader, forward-looking view of how Cartwright's research could continue to shape evidence-based policymaking in healthcare and other sectors.
Rater 3 Rating = 4 (Very Good)
IR Justification: The ChatGPT content provided insightful implications, especially in discussing future potential shifts in empirical methodology and the broad adoption of alternatives to RCTs. However, its predictions on policy impact and educational reforms did not capture the same level of detailed stakeholder influence described in the REF-submission.
Rater 4 Rating = 5 (Excellent)
IR Justification: The ChatGPT-generated impacts provide insightful projections on how Cartwright's critiques of RCTs could continue shaping policy methods, especially by promoting alternative approaches that consider complex social and economic factors. This forward-looking aspect aligns well with the REF impacts and expands on Cartwright's broader theoretical influence. The ChatGPT version also extrapolates potential long-term shifts in policy-making

globally, a factor only indirectly suggested in the REF submission.

"
"27. Study Title : Seeing beyond the wheelchair: Pioneering education and higher aspiration promotion for boys and men with Duchenne Muscular Dystrophy
Rater 1 Rating = 4 (Very Good)
IR Justification: ChatGPT offers insightful forward-looking predictions, especially regarding applicability to other neurodevelopmental conditions.
Rater 2 Rating = 4 (Very Good)
IR Justification: The AI-generated impacts demonstrate significant insight into the potential broader effects of the research, extending the implications to future educational policies, and healthcare reforms. This adds layers of interpretation that go beyond the immediate impacts discussed in the REF submission. The REF case study, while thorough, is more focused on describing realized impacts, without as much forward-looking projection.
Rater 3 Rating = 4
IR Justification: AI output exhibits a high level of insightfulness by projecting additional applications of the research for broader neurodevelopmental conditions and community-based education. Additionally, it emphasizes the enduring relevance of advocacy and educational materials beyond DMD, which aligns with future impact considerations. However, it sometimes lacks precise grounding in existing policy measures, which limits the depth of actionable insights provided.
Rater 4 Rating = 4 (Very Good)
IR Justification: The ChatGPT version offers insightful discussions on the social and educational implications of early intervention for DMD, highlighting future potential applications in other educational policies for genetic disorders. This level of insight suggests forward-thinking analysis but lacks certain specific policy recommendations noted in the REF submission.

"
"28. Study Title : The Coffee Historian: Achieving Impact Through Industry Collaboration, Education, and Public Engagement
Rater 1 Rating = 3 (Good)
IR Justification: While the ChatGPT version provides insightful observations, including predicting potential future impacts, such as the commercialization of espresso machine innovations, it lacks the depth of specific case studies provided in the REF submission. The REF version's detailed descriptions of industry impact provide a clearer understanding of real-world applications.
Rater 2 Rating = 3 - Good
IR Justification: While ChatGPT-generated impacts provide a strong overview, they lack some of the deeper insights found in the REF version. The REF case study highlights how Morris's work specifically influenced product narratives and cultural perceptions, which is less predictive and forward-thinking in the ChatGPT version. However, ChatGPT's ability to frame the research in terms of future trends in the coffee industry adds a layer of insightfulness in predicting technological shifts.
Rater 3 Rating = 3 (Good)
IR Justification: The ChatGPT version presents insightful extrapolations, especially around future impacts and scalability, offering suggestions for additional international collaborations and policy pathways. It misses, however, some of the nuanced insights on Italian coffee culture's evolution and the direct consumer impacts found in the REF version.
Rater 4 Rating = 4 (Very Good)
IR Justification: The AI version demonstrates insight into potential industry impacts by analyzing scalability, future consumer preferences, and the adaptability of Italian coffee models, aligning well with academic and commercial interests. However, while it predicts future applications and cross-sector relevance, it lacks some of the nuanced industry-specific insights that the REF submission addresses through real-world applications.

"
"29. Study Title : Discovering Ted Hughes's Yorkshire
Rater 1 Rating = 5 (Excellent)
IR Justification: The AI-generated version offers valuable insights into how Hughes's work could be understood through a lens of psychological and ecological healing, which extends beyond the literal cultural heritage impacts covered in the REF submission. It also touches upon broader academic fields, adding predictive insights about the future importance of these themes in different disciplines.
Rater 2 Rating = 3 (Good)
IR Justification: While the ChatGPT version provides insightful commentary on potential impacts,

especially in terms of future policy and societal shifts, it lacks the predictive depth and detailed future projections related to specific community benefits seen in the REF submission. The AI version does not expand on how specific programs like the heritage trails might evolve or how educational initiatives might shape future cultural landscapes.
Rater 3 Rating = 3
IR Justification: The AI provides an insightful overview of the impacts but lacks some of the contextual depth present in the REF submission, such as the nuanced role of individual contributors and the specific heritage-enhancing actions taken. However, it adds value by predicting wider societal and ecological implications of increased cultural tourism.
Rater 4 Rating = 3
IR Justification: The ChatGPT version offers broad insight, suggesting possible regional scalability and relevance to cultural tourism. However, it lacks some of the granular predictive insights about long-term cultural and economic regeneration effects, as well as emerging community benefits from place-making activities uniquely present in the REF.

"
"30. Study Title : Empowering Indigenous Self-Representation for the Emberá People of Panama
Rater 1 Rating = 3 (Good)
IR Justification: The AI-generated content provides meaningful insight into the broader implications of the research, particularly in terms of sustainability and the empowerment of indigenous communities. However, the REF submission offers a more nuanced understanding of the political implications and the direct use of research in advocacy and legal processes, which adds depth.
Rater 2 Rating = 4: Very Good
IR Justification: The AI-generated report provided valuable insights into future trends and potential applications of the research findings beyond the immediate context. It predicted the long-term socio-political impacts on indigenous communities globally and offered foresight into how modernity and tradition might evolve in tourism practices. These elements add significant value beyond the direct impacts outlined in the REF submission.
Rater 3 Rating = 4: Very Good
IR Justification: The AI-generated insights reflect a high level of understanding by emphasizing future interdisciplinary collaborations and the scalable impact of Indigenous tourism models. This interpretation aligns with the REF's focus on the broader relevance of this research but adds foresight into potential international and policy-level applications. However, the AI version lacks some of the personalized insights seen in community testimonials from the REF submission, which provide additional depth on the social impacts of this project within the Emberá community.
Rater 4 Rating = [4] Very Good
IR Justification: The AI-generated impacts provide insight into future implications of cultural preservation through self-representation, suggesting sustained benefits. REF submission focuses on specific legal and sociocultural gains but with less projection of future implications, marking the AI content as insightful though requiring more regional applicability.

"
"31. Study Title : Enhancing Understanding of the Foreign and Security Policy Implications of Brexit for Government and the Wider Policy Community
Rater 1 Rating = 4 (Very Good)
IR Justification: ChatGPT adds insight by providing predictions about the long-term implications of Brexit on global trade and EU diplomatic strategies. The REF version focuses more on immediate, tangible impacts on policy, but the ChatGPT version extends these into possible future scenarios, providing useful forward-looking analysis.
Rater 2 Rating = 3 - Good
IR Justification: The AI version provides some forward-looking insights, particularly in its discussion of the long-term geopolitical consequences of Brexit. However, while it offers good potential projections, it lacks depth in predictive insights regarding specific policy shifts and stakeholder reactions, which could be derived from the detailed interactions noted in the REF version.
Rater 3 Rating = 4
IR Justification: The AI-generated impacts provide thoughtful projections of the research's influence on EU policy and long-term UK diplomacy, which is insightful in recognizing broader geopolitical consequences. The AI content offers meaningful extensions, especially in discussing future implications for the EU's foreign policy strategies with non-member states, which is slightly beyond the REF submission's more UK-focused scope.
Rater 4 Rating = 5 (Excellent)
IR Justification: The AI-generated impacts provide significant insights, projecting the UK's future diplomatic positioning in light of Brexit, including adaptability within evolving global

security frameworks. It introduces future-oriented scenarios that encourage proactive defense alignment.

"

"32. Study Title : Worldwide Improvements in Policing due to Increased Sales of Facial Composite Software
Rater 1 Rating = 4 (Very Good)
IR Justification: The ChatGPT version provides a good level of insight, particularly in its prediction of future trends in technological transferability and long-term applications in fields like healthcare and entertainment. This forward-looking perspective provides additional insights not explicitly mentioned in the REF version, offering a broader vision of the technology's potential.
Rater 2 Rating = 4
IR Justification: The AI-generated impacts show insight into future trends and broader technological implications, extending beyond the immediate scope of the facial composite software's impact on policing. This forward-thinking approach adds value, as it explores potential long-term benefits in other domains.
Rater 3 Rating = 3 (Good)
IR Justification: The AI-generated version provides a forward-looking view, considering future technological adaptations across fields, offering a more expansive vision. However, its insights on practical law enforcement applications and current procedural efficiency—covered thoroughly in the REF report—are less detailed in the AI-generated content.
Rater 4 Rating = 4: Very Good
IR Justification: The AI version is insightful in proposing future impacts and suggesting broader applications, such as in healthcare and digital identity sectors. However, it lacks certain predictive depth about specific changes in policing methodologies.

"

"33. Study Title : Improving the Physical Wellbeing of the Police Force
Rater 1 Rating = 3 (Good)
IR Justification: The AI-generated content provides insightful reflections on broader implications and scalability but lacks some forward-looking analysis seen in the REF version. It also could have gone deeper into emerging trends based on the existing program's success.
Rater 2 Rating = 3: Good
IR Justification: The AI-generated content offers insights into the broader public health implications of the Fitness Mentors Program, which could lead to new applications beyond the police force. However, it lacks some of the more immediate, context-specific insights present in the REF submission, such as participant testimonials and the direct impacts on local policy.
Rater 3 Rating = 4
IR Justification: The AI version provides meaningful insights into how the research could influence public safety and wellbeing on a broader scale, suggesting implications for other emergency services and public sector wellbeing programs. However, it lacks the in-depth future outlook specific to Lincolnshire Police's further training plans and adaptation in response to COVID-19.
Rater 4 Rating = 5 (Excellent)
IR Justification: The AI-generated report provides insightful considerations regarding potential future directions for similar wellness initiatives in other sectors, drawing connections beyond the immediate outcomes. It offers possible policy implications and suggestions for adaptation in different fields. The AI analysis suggests enduring impacts that the REF submission did not cover as explicitly, enhancing the report's foresight.

"

"34. Study Title : Raising Maori students' achievement in secondary schools in Aotearoa New Zealand
Rater 1 Rating = 4
IR Justification: The AI-generated content includes forward-looking insights, particularly about how the educational methods studied could impact future curriculum design and policy frameworks. The REF version, while rich in data and present outcomes, does not emphasize as much on future potential beyond the specific Māori context.
Rater 2 Rating = 4 (Very Good)
IR Justification: The AI-generated version provides insightful observations, especially regarding the potential for the educational interventions to be applied in different contexts and their long-term impacts on educational policy. This perspective on future impacts provides additional value not emphasized in the REF submission, which is more focused on immediate, measurable outcomes. The AI-generated impacts also mention trends in literacy and inclusive pedagogy that could influence future educational reforms.

Rater 3 Rating = 3 (Good)
IR Justification: The ChatGPT version provided insightful interpretations of the research's impact, particularly in terms of broader implications and future applications. However, it could further enhance its insightfulness by predicting specific potential shifts in policy or teacher training as a result of these impacts.
Rater 4 Rating = 4 (Very Good)
IR Justification: The AI-generated impact includes forward-looking insights into possible applications of culturally responsive pedagogy in other contexts, such as similar frameworks for other indigenous groups globally, showing a well-grounded understanding of the study's future potential. However, certain immediate application insights, like the specific observed changes in student retention and engagement, are less prominent than in the REF submission, which slightly limits its depth of immediate insight.

"
"35. Study Title : Improving patient outcomes and treatment guidelines through the study of Hepatitis C
Rater 1 Rating = 4 (Very Good)
IR Justification: The ChatGPT-generated content is insightful in predicting long-term implications and global transferability. It identifies areas for future research and technological development (e.g., automated referral systems), which extends beyond the direct impacts mentioned in the REF submission.
Rater 2 Rating = 4 (Very Good)
IR Justification: The AI-generated impacts offer insightful reflections on potential future impacts and broader implications of the research, particularly in terms of its scalability to other health challenges and global adoption of DAAs. This level of insight adds value beyond what is directly stated in the REF submission, predicting further global trends in disease management. However, it lacks some context-specific insights that come from the testimonies and policy nuances present in the original submission.
Rater 3 Rating = 3
IR Justification: The AI version is moderately insightful, speculating on future policy adaptations and scalability of healthcare improvements but lacking in-depth actionable recommendations found in the REF. It projects potential impacts without fully capturing the long-term frameworks or direct clinical guidelines highlighted in the REF.
Rater 4 Rating = 3 - Good
IR Justification: The AI-generated content provides useful insight into potential long-term health impacts and suggests broader applications, especially in healthcare policy and patient outcomes. Yet, it lacks some of the in-depth projections present in the REF submission, such as predictive models for DAA impacts on patient longevity and specific epidemiological data supporting the cascade of care intervention. The AI content also could benefit from more detailed discussion on the potential impacts of resistance tracking and implications for broader HCV treatment frameworks.

"
"36. Study Title : Improving Homecare Quality in the UK Through Optimized Workforce Planning
Rater 1 Rating = 4 - Very Good
IR Justification: The AI-generated report provides insightful extensions, such as predicting potential future applications of the workforce optimization algorithms in new sectors like urban planning and healthcare, which extend beyond the original scope of the REF submission. It also considers the global relevance of the research.
Rater 2 Rating = 3 (Good)
IR Justification: Insightful, with forward-looking perspectives but lacking detailed real-world evidence.
Rater 3 Rating = 3 (Good)
IR Justification: The AI-generated analysis offered insightful extensions, including potential future impacts and scalability across industries, yet did not reach the depth of operational specifics that the REF submission covered in workforce satisfaction metrics and client growth figures.
Rater 4 Rating = 5 (Excellent)
IR Justification: The AI-generated impacts provided a high level of insight, especially in discussing the scalability and long-term industry impacts of OptifAI. It highlighted potential broader applications, future economic impacts, and the disruptive potential of scheduling optimizations in related sectors. This extrapolative insight added depth to the original REF submission.

"
"37. Study Title : Transforming vaccine policy for pneumococcal disease leading to significant

cost savings in the NHS
Rater 1 Rating = 4 - Very Good
IR Justification: The ChatGPT version offers insightful projections on the broader implications
of the research, including potential technological disruptions and applications in other
healthcare diagnostics. It also forecasts how the assay could influence long-term vaccine policy
and public health strategies globally. The REF version, while more detailed, does not speculate
as much on future impacts outside of the immediate research applications.
Rater 2 Rating = 3: Good
IR Justification: AI provides some forward-looking predictions about future policy trends,
though it lacks the detailed, evidence-based policy reflections seen in the REF submission.
Rater 3 Rating = 5: Excellent
IR Justification: ChatGPT provided insightful predictions on future applications of the research
in health surveillance and diagnostics, highlighting advancements in non-invasive diagnostic
applications and implications for policy evolution, which were not present in the REF submission.
Rater 4 Rating = 4
IR Justification: The AI-generated content provides forward-looking insights, particularly on
the potential adaptation of research outcomes in broader health contexts. It suggests meaningful
applications, extending insights beyond those in the REF submission, which could potentially
enrich the scope of the impact narrative. However, some forecasted impacts could benefit from
concrete examples of application as provided in the REF submission.

"
"38. Study Title : UoP32Househistories: A House Through Time: Shaping a flagship TV series to
achieve critical and financial success and inspiring the public to engage with house history
Rater 1 Rating = 4 (Very Good)
IR Justification: The ChatGPT-generated content is insightful, particularly in its ability to
predict the long-term cultural and policy impacts of the research. It discusses the potential
future trends in public engagement with heritage conservation and how the series can influence
policy decisions in urban planning and historical preservation. However, the insights are not as
specific to the study's practical contributions to media production, which were clearer in the
REF submission.
Rater 2 Rating = 3 (Good)
IR Justification: While the AI-generated impact provides some insightful connections,
particularly in anticipating future developments based on the program's success, it does not
significantly extend beyond the content of the REF submission. The REF version offers richer
context, such as how public interest in house histories grew from a scholarly topic to broader
engagement, which the AI version could build upon.
Rater 3 Rating = 4 (Very Good)
IR Justification: The ChatGPT version demonstrates insight into potential public engagement
impacts and the historical significance of suburban and domestic history. The REF submission,
however, includes further details on Ryan's expert contributions to visual flair and historical
accuracy, adding depth to its insights.
Rater 4 Rating = 4 (Very Good)
IR Justification: The AI-generated report is insightful in highlighting how the program might
affect public interest in historical and material culture studies, going beyond the immediate
impact described in the REF submission. This prediction of a sustained cultural shift aligns
well with the REF's documented audience response, albeit with less specificity regarding
particular archival or academic impacts.

"
"39. Study Title : Optimising baggage operations at London Heathrow Airport to achieve cost
savings for the aviation industry
Rater 1 Rating = 4 (Very Good)
IR Justification: The AI-generated version provides significant insights into the broader
implications of the research, particularly regarding potential future applications in various
fields. The REF submission, on the other hand, is more pragmatic, detailing current impacts and
measured improvements in operational efficiency at Heathrow. The ChatGPT version adds value by
suggesting new avenues for research and implementation, although some of these insights are not
directly tied to the immediate case study.
Rater 2 Rating = 3 (Good)
IR Justification: While the AI-generated impacts are meaningful and contain actionable
information, they do not extend much beyond the original case study's stated impacts. The REF
submission included more forward-looking insights, such as plans for further implementation
(e.g., expansion to Terminal 5) and the potential for addressing post-COVID operational changes.
The AI version did not delve into these forward-looking scenarios as deeply.
Rater 3 Rating = 4 (Very Good)

IR Justification: The AI-generated content offered insight into future applications and implications of the research, predicting potential use cases across similar facilities. This perspective goes beyond the immediate impacts, aligning with insightful aspects seen in comprehensive impact assessments.
Rater 4 Rating = 4 - Very Good
IR Justification: The AI report demonstrated insightful predictions about how optimisation methods could continue to impact operational efficiency and passenger experience, suggesting future applications in similar contexts. It made meaningful suggestions on the resilience of dynamic scheduling in adapting to unpredictable events, which adds depth.

"
"40. Study Title : POWeR - Cost-effective online support for weight management
Rater 1 Rating = 4
IR Justification: The ChatGPT version provides insightful analysis by predicting the potential for long-term scalability of the digital intervention and offering suggestions for broader technological integration. It also considers implications for future healthcare innovations. However, some of the insights are more speculative compared to the evidence-based focus of the REF submission.
Rater 2 Rating = 3: Good
IR Justification: While the AI-generated content is solid, it largely restates the established impacts without adding significant novel insights or analysis beyond what is in the original REF submission. It provides some foresight into potential future applications of POWeR in different contexts, which is valuable, but it does not delve deeply into broader implications or predictive trends about the project's evolving impact over time. This contrasts with the original submission, which was more focused on current impacts and specific stakeholder feedback.
Rater 3 Rating = 4 (Very Good)
IR Justification: The AI-generated text offers insightful predictions about potential implications, such as reduced healthcare resource demands and increased accessibility for low-cost weight management solutions. It also highlights the possible long-term benefits in healthcare sustainability, though lacks specific examples of ongoing international applications noted in the REF.
Rater 4 Rating = 4
IR Justification: The ChatGPT version demonstrates foresight into how the POWeR model can influence future digital healthcare initiatives, particularly for chronic disease management. This expands upon the REF's focus on obesity and primary care settings, reflecting well on its insightfulness. However, the REF submission provides more actionable insights by detailing contractual obligations, which are critical for practical applications.

"
"41. Study Title : The global impact of Sunderland's football research on scouting, training, and player preparation in elite football
Rater 1 Rating = 3 - Good
IR Justification: The AI-generated content is insightful in its extrapolation of the research impacts, particularly in suggesting broader implications like injury prevention and health optimization. However, it does not delve as deeply into the long-term influence on sports policy and stakeholder engagement, which the REF submission addresses with more specificity.
Rater 2 Rating = 3: Good
IR Justification: While the AI-generated version presents meaningful insights, particularly about future technological applications and potential cross-disciplinary impacts, it does not go as deeply into actionable insights that would drive further research or policy change. The REF submission is more direct in illustrating how the research shaped training regimens and injury prevention strategies, providing specific examples of adaptation in practice.
Rater 3 Rating = 4: Very Good
IR Justification: The AI-generated content provides insightful extrapolations about the influence of Sunderland's research on player preparation and resilience. It offers forward-looking statements on future impact on coaching tech and player management beyond what is explicitly stated in the REF submission. However, some implications are general and do not fully utilize specific case examples from the REF submission, which contain detailed narratives from influential sports figures.
Rater 4 Rating = 4 (Very Good)
IR Justification: ChatGPT content introduces insightful considerations, particularly regarding the potential for evolving analytics to adapt training techniques in the future. It also anticipates the role of these findings in predictive injury prevention models, offering an advanced perspective. However, it lacks explicit predictions about future studies in non-European football leagues, which the REF content subtly alludes to.

"42. Study Title : Adding value to convenience retailing through improved pricing, a new store concept, and capacity building
Rater 1 Rating = 3 (Good)
IR Justification: While the AI report provides a good understanding of the immediate and transferable impacts of the research, it lacks some depth in predicting long-term societal trends or providing detailed actionable insights for practitioners. The REF submission offers richer practical insights directly from stakeholder feedback.
Rater 2 Rating = 3 (Good)
IR Justification: The AI-generated content provides good insights into the future implications of the research, such as predictions on the scalability of the store concept and potential for influencing future retail trends. However, the insights were not as specific as the REF submission, which provided more detailed and practical examples, like workshops and direct contributions to award-winning strategies.
Rater 3 Rating = 4 (Very Good)
IR Justification: The AI-generated report provides insightful analyses of future implications and potential scalability of the retail convenience model. It extends beyond the immediate results to suggest broader impacts on retail technology adoption and consumer satisfaction trends. This additional insight helps predict possible future directions for the retail industry, though it lacks the specific depth of industry-focused testimonies provided in the REF submission.
Rater 4 Rating = 4
IR Justification: ChatGPT provides insights into broader implications, such as future shifts in consumer demographics and store layout trends in retail, extending the original study's implications into potential future impacts. However, it could deepen on aspects like sustainability impacts of the new concepts, which are only tangentially referenced.

"
"43. Study Title : Transforming Evidence-Based Practice in Public Health Through Co-Production and Evaluation
Rater 1 Rating = 3 (Good)
IR Justification: While the AI-generated impacts offered insightful commentary on the scalability of co-production research and future applicability, it tended to reiterate known outcomes rather than offer forward-looking insights or predictions. The AI content was strong in terms of methodological insight but less robust in predicting future trends or broad policy shifts.
Rater 2 Rating = 4 (Very Good)
IR Justification: The AI version offers insightful projections about how the co-production model could influence broader public health interventions. It goes beyond the immediate outcomes discussed in the REF submission by anticipating future uses of co-production frameworks in policy and education settings. However, some of these insights lack specificity compared to the detailed evaluations and their documented impact on policy changes present in the REF case studies.
Rater 3 Rating = 5 (Excellent)
IR Justification: The AI-generated content demonstrates a high level of insightfulness by extending the co-production model to predict broader societal impacts and offering insights on the ethical and sustainable use of co-produced evidence. This adds depth to the practical implications of co-production beyond immediate effects.
Rater 4 Rating = 4
IR Justification: The AI-generated content provides insightful extensions, such as the scalability of the co-production model and broader implications for international health systems. However, while it discusses future applicability, it does not predict or analyze specific community outcomes or changes stemming from the sustained use of these models as the REF does.

"
"44. Study Title : Well-being and Public Policy
Rater 1 Rating = 4 (Very Good)
IR Justification: The AI-generated impacts provide insightful predictions about the future of well-being metrics, particularly regarding the long-term adoption of community-based health interventions. It offers foresight into potential new research avenues and public policy developments. However, the AI content is less specific in describing existing stakeholder collaborations and how past interventions influenced current policy. The REF submission, on the other hand, provides detailed examples of past actions that shaped policy, offering a more grounded insight.
Rater 2 Rating = 5 (Excellent)

IR Justification: The AI-generated report demonstrates strong insightfulness, providing future-oriented interpretations of the research impacts. It extends beyond merely repeating what was done and offers plausible predictions for how the well-being research could influence long-term policy directions, particularly in areas of well-being measurement and social interventions. This future-looking perspective adds value to the evaluation. Additionally, it reflects on potential applications in other countries and contexts, highlighting the scalability and adaptability of the findings.
Rater 3 Rating = 4
IR Justification: The AI-generated content provides meaningful insights and contextualizes potential implications for well-being metrics in public policy. However, the REF submission includes a nuanced discussion of how these metrics shape cross-departmental government initiatives and policy decisions, which offers additional depth beyond the ChatGPT version.
Rater 4 Rating = 4 (Very Good)
IR Justification: The AI-generated impacts introduced potential future applications of well-being research, extending insights on public policy to international bodies and anticipating more substantial outcomes in societal well-being metrics. This predictive insight was valuable, although it occasionally extrapolated beyond the scope presented in the REF document. Insights on societal trends and potential future applicability added to the AI version's depth, albeit with some speculative elements.

"
"45. Study Title : Improving the quality of green infrastructure in towns and cities in the UK
Rater 1 Rating = 3 - Good
IR Justification: The AI-generated version provides valuable insights into the broad application and future trends of green infrastructure, particularly around the global relevance of the research. However, the REF submission provides deeper insight into the direct policy influence and the strategic long-term partnerships with stakeholders like Building with Nature. The AI version lacks the same depth of insight into long-term local policy evolution.
Rater 2 Rating = 3
IR Justification: While the ChatGPT content reflects a reasonable degree of insight, it does not extend beyond what was explicitly stated in the REF submission. Predictions about long-term impacts and trends could have been further developed. The REF content included more nuanced projections regarding urban development and future challenges, which would have added depth to the AI's version.
Rater 3 Rating = 4 (Very Good)
IR Justification: ChatGPT's generated content includes insights into long-term potential impacts, like possible policy changes and the emphasis on stakeholder engagement in green infrastructure design. However, the AI-generated version lacks some of the detailed predictions found in the REF submission regarding stakeholder-specific changes and exact projected economic impacts.
Rater 4 Rating = 5: Excellent
IR Justification: The AI-generated impact extends beyond the case study, suggesting future implications and adaptability of green infrastructure standards to other regions and potential climate-related challenges, which adds valuable predictive insights. This proactive analysis adds a layer of foresight beyond the original REF submission, contributing to policy evolution and infrastructure scalability.

"
"46. Study Title : Redressing the state of the stateless: seeking political recognition for Tibet and Kashmir
Rater 1 Rating = 3: Good
IR Justification: The AI-generated content provides meaningful insights into how the research may extend into broader academic and policy contexts, especially regarding cultural representations and identity politics. However, the REF version demonstrates a more grounded insight into the real-world challenges faced by the researchers, including personal and social obstacles, which makes the REF version slightly more insightful regarding the researchers' personal involvement.
Rater 2 Rating = 3 - Good
IR Justification: While insightful, the AI-generated version is more descriptive than predictive or critical. It captures the main impacts but lacks depth in predicting future trends or the broader implications of these impacts on international policy and academia. Insightfulness is achieved to a reasonable extent but could benefit from more forward-looking analysis.
Rater 3 Rating = 4
IR Justification: The AI version provides valuable insights by projecting the long-term impacts on public policy and societal perception of stateless populations, demonstrating understanding of the advocacy trajectory for Tibet and Kashmir beyond present impacts. The AI version adds

potential future relevance in postcolonial studies, furthering the discussion in ways that can foster deeper academic interest.
Rater 4 Rating = 3 (Good)
IR Justification: While the AI-generated content draws meaningful connections between research and its broad social influence, it does not provide predictive insights into future impacts of advocacy or legislative changes, as detailed in the REF submission's analysis of long-term case law influence.

"
"47. Study Title : Evidence-based enteral feeding practices for very preterm or very low birth weight infants
Rater 1 Rating = 4 (Very Good)
IR Justification: The AI-generated content offers predictions on long-term impacts and provides broader insights, such as how the research could influence future technological innovations in nutrition. It goes beyond just reporting current impacts and suggests future research pathways and potential global policy shifts.
Rater 2 Rating = 4 (Very Good)
IR Justification: The ChatGPT-generated content demonstrates insightfulness by projecting how the research could influence future studies and health policies globally. It anticipates potential advancements in enteral feeding practices and economic benefits through cost-saving measures. The REF submission focuses more on the current state of research, but the ChatGPT version extends the discussion toward future research opportunities and trends.
Rater 3 Rating = 4: Very Good
IR Justification: The ChatGPT-generated content provides additional insights on long-term policy influence and potential curriculum integration, showcasing depth in implementation pathways. Predictions for broader applications in health systems and the public sector add thoughtful foresight, although it lacks some precise impact examples provided in the REF on patient outcomes like infection reduction.
Rater 4 Rating = 3
IR Justification: The AI version offers insightful predictions on how feeding interventions may continue to shape neonatal practices globally. However, it lacks foresight on specific upcoming challenges in policy adoption or future barriers to implementation that the REF submission partially hints at through detailed international data.

"
"48. Study Title : Improving Environmental Conservation in East Africa and Beyond
Rater 1 Rating = 3 (Good)
IR Justification: ChatGPT's insights are generally good but not always as forward-looking or predictive as the REF submission. While it captures the immediate impacts and offers some long-term predictions, it does not always delve into the specific future trends or policy implications as deeply as the REF submission.
Rater 2 Rating = 3
IR Justification: The AI-generated report provides insightful analyses regarding future technological applications and broader trends in conservation management, such as integrating ranger-based data collection for optimized patrols. However, it does not venture as deeply into predictive impacts on socio-economic structures or the policy-related effects that the original research underscores, such as its impact on the IUCN Red List and community conservation agreements.
Rater 3 Rating = 4 (Very Good)
IR Justification: The AI-generated version introduces some insightful additions, such as potential scalability to global conservation efforts and technological innovation via conservation tools like SMART. However, it does not fully capture the nuances of policy influence as comprehensively as the REF submission.
Rater 4 Rating = 4: Very Good
IR Justification: The ChatGPT-generated impacts add insightful commentary, projecting potential global applications and additional impacts within conservation fields. This insightfulness includes anticipated advancements in ecological restoration techniques, with predictions on long-term outcomes and alternative uses of AFRICLIM. However, while innovative, some insights lack the empirical grounding present in the REF data.

"
"49. Study Title : Predicting the properties of materials with first-principles electronic structure software (CASTEP)
Rater 1 Rating = 4 (Very Good)
IR Justification: The AI version provides some forward-looking insights, predicting future trends in energy-efficient electronics and broader technological disruptions, which enhances the

depth of the analysis. The REF submission, while highly detailed, is more grounded in present and past achievements, whereas the AI version speculates more on future innovations.
Rater 2 Rating = 4 (Very Good)
IR Justification: The AI-generated content provided some novel insights, such as future implications for quantum computing and energy-efficient electronics, which were not fully explored in the REF submission. This shows an understanding of the broader, longer-term impacts of the technology. However, these insights are speculative and would benefit from further validation or supporting evidence.
Rater 3 Rating = 4
IR Justification: The AI-generated impacts provide forward-looking insights, including potential for CASTEP to influence green technologies and broader applications in biomedicine. This provides actionable directions that extend beyond the REF submission, though these insights are speculative rather than substantiated by current CASTEP applications.
Rater 4 Rating = 3 - Good
IR Justification: While the AI-generated version captured meaningful insights into CASTEP's technological advancement, it provided limited predictions for the software's future developments or potential shifts in industrial use. It could have extended analysis on emerging trends in electronic structure research or anticipated advancements in CASTEP's functionality.

"
"50. Study Title : Statistical pattern recognition applied to protein crystallisation images in the pharmaceutical industry
Rater 1 Rating = 4 - Very Good
IR Justification: The AI-generated impact analysis provided valuable insights into the long-term implications of the research, especially in terms of scaling automation in crystallography and transferring methods to other fields like genomics. It also discussed broader societal implications, such as sustainability, which the REF submission did not focus on.
Rater 2 Rating = 4
IR Justification: The AI-generated impacts offer significant insight by predicting future directions of the research, such as its scalability to other industries and the potential for disruptive technology across fields like biopharmaceuticals. However, the insights remain somewhat general compared to the specific and practical examples in the REF submission (e.g., the exact role of MARCO in accelerating crystallization processes in various companies).
Rater 3 Rating = 4: Very Good
IR Justification: The AI-generated impacts demonstrate a high level of insight, addressing both immediate and long-term implications, such as the expansion of automation to other high-throughput labs and the potential reduction in labor costs. It also anticipates possible future applications in other areas of drug discovery and structural biology. This demonstrates a forward-looking perspective that complements the REF's technical depth.
Rater 4 Rating = 4 (Very Good)
IR Justification: The AI-generated version provides forward-looking insights on the scalability and adaptability of the research outcomes across industries and future healthcare applications. It effectively projects future advancements in automation in similar experimental contexts and suggests broader applications of AI in laboratory automation.

"
"51. Study Title : 'Moving beyond one-size-fits-all: Improving Widening Participation through Realist Evaluation methodologies in Northern England'
Rater 1 Rating = 4
IR Justification: The AI-generated version provides insightful analysis by predicting broader societal impacts of the research, such as the potential for policy changes at national levels. It also extrapolates the findings to global education systems, which is not explicitly mentioned in the REF submission. However, the REF submission provides more granular, practice-based insights directly connected to the outcomes of the program.
Rater 2 Rating = 3 (Good)
IR Justification: While the AI-generated content is insightful in terms of recognizing the potential long-term implications of using RE in WP, it doesn't go as far as the REF submission in anticipating future outcomes based on current evidence. The REF version provides detailed reflections on how interventions can be fine-tuned to local contexts, offering deeper insights into why particular strategies work in specific settings.
Rater 3 Rating = 3
IR Justification: The ChatGPT version gives meaningful interpretations and anticipates possible policy extensions, demonstrating forward-thinking beyond the explicit impacts in the REF version. However, it does not fully capture some implicit insights into the cultural shifts due to WP initiatives, as detailed in the REF submission.
Rater 4 Rating = 3 (Good)

IR Justification: While the AI-generated impacts were insightful and captured foundational implications for community outreach and HE accessibility, the analysis did not fully extend to predictions or strategic recommendations that could add a forward-looking perspective. The original REF submission provided greater context regarding the unique challenges in implementing RE within WP, which added a deeper layer of reflection on the practical implications of RE methodologies.

"

"52. Study Title : Challenging monolithic conceptualisations of English for learning, teaching and assessment: The Changing Englishes online course
Rater 1 Rating = 4 - Very Good
IR Justification: The AI-generated content demonstrates strong insight into the potential for future policy shifts, the disruption of traditional language teaching frameworks, and the importance of cognitive and sociolinguistic diversity in education. It suggests actionable predictions on how this research could shape future trends.
Rater 2 Rating = 4 (Very Good)
IR Justification: The AI-generated impacts were insightful in predicting long-term effects, such as influencing international language policies and creating innovations in language testing frameworks. This forward-thinking analysis extended beyond the REF's immediate focus on pedagogical change. However, some deeper societal implications, like the impact on linguistic identity and mental health, were only briefly touched on.
Rater 3 Rating = 3 (Good)
IR Justification: The AI-generated impacts bring forward ideas on future policy trends and the wider application of plurilithic approaches, which adds value beyond what is explicitly stated in the REF submission. However, the REF submission presents more insight into the specific transformative effects on teacher mindsets and classroom applications.
Rater 4 Rating = 3 - Good
IR Justification: While the AI version provides valuable insights into future applications and the potential for expanded frameworks in multilingual settings, it lacks some of the REF's detailed, firsthand evidence of transformative experiences for educators. Consequently, the insight is meaningful but lacks some depth of practical nuance.

"

"53. Study Title : Robert Paul and the Origins of British Cinema
Rater 1 Rating = 3 (Good)
IR Justification: ChatGPT's version provides meaningful insights into how technological advancements might influence future trends in advertising and cultural heritage preservation. However, it does not offer as many actionable insights into community-level impacts or long-term cultural engagement as seen in the REF submission.
Rater 2 Rating = 3 - Good
IR Justification: While the AI-generated impact identifies key trends in technological and cultural impacts, it lacks deep insight into the long-term historical and societal significance of Paul's contributions beyond the basic preservation of film. The REF version adds context on how these films influenced public awareness of early cinema and historical understanding, particularly through community engagement efforts.
Rater 3 Rating = 5 (Excellent)
IR Justification: ChatGPT's generated impacts provide meaningful insights, especially in projecting the influence of early cinematic technology on modern media and advertising. The analysis hints at the long-term cultural significance of Paul's work and its ongoing relevance, showing depth beyond the case study's explicit content.
Rater 4 Rating = 4: Very Good
IR Justification: The AI version includes insightful analysis by exploring potential future applications of Paul's work, such as advancements in film preservation and increased global engagement with early British cinema. This future-oriented insight is valuable, though the REF submission provides stronger insights into the community impact of exhibitions and the actual responses from museum visitors, as well as the tangible outcomes of these efforts on local institutions.

"

"54. Study Title : The Care and Management of Gout in Primary Care
Rater 1 Rating = 3 (Good)
IR Justification: The AI-generated version does provide actionable insights, such as future applications for digital tools and potential improvements in treatment adherence. However, it falls short of the REF version's depth in capturing the direct consequences of the research on specific clinical practices, like the nuances of early pharmacological intervention for gout and the impact on kidney disease management.

Rater 2 Rating = 4: Very Good
IR Justification: The AI-generated content provides meaningful extensions of the REF impacts, including potential future health interventions and policy changes, particularly in regions with higher gout prevalence. It also introduces some novel perspectives, such as the long-term economic benefits of improved gout management strategies. However, while insightful, some of these predictions remain speculative, and the AI lacks the depth found in the qualitative insights of the original case study regarding patient experience.
Rater 3 Rating = 4
IR Justification: The AI-generated content demonstrated good insight, particularly regarding future applications in clinical training and the sustainability of patient education resources. It anticipated policy applications beyond the immediate scope but missed some actionable elements from the REF version on nurse-led care.
Rater 4 Rating = 5 (Excellent)
IR Justification: The AI-generated version presents insightful interpretations of the REF findings, providing actionable recommendations for future research directions and recognizing the emerging importance of integrated care for gout in broader comorbid contexts, such as cardiovascular and kidney health. This goes beyond the initial REF focus by predicting downstream impacts on other chronic disease management.

"
"55. Study Title : From Victims to Actors: Shifting the Policy Paradigm to Value Children's Contribution in Disaster Risk Management
Rater 1 Rating = 4 (Very Good)
IR Justification: The AI-generated impacts were quite insightful, providing forward-looking analysis on the potential for policy influence and suggesting practical tools for broader dissemination, such as apps or educational programs. These insights extend the implications of the research beyond what was explicitly stated in the REF submission.
Rater 2 Rating = 4 - Very Good
IR Justification: The AI-generated impacts demonstrated solid insight into how the research could extend beyond the immediate project. For example, the AI-generated content explored potential future policy reforms based on the findings and suggested trends for integrating children's voices more formally in disaster risk management. This predictive nature adds a valuable dimension that was less pronounced in the REF version, which focused more on reporting already achieved impacts.
Rater 3 Rating = 3 - Good
IR Justification: The AI report offers some predictive insights, such as the long-term potential of integrating children's roles in disaster policies globally. It lacks in-depth predictions on potential future UK policy changes or direct quotes to underline the emotional and practical effects on child participants highlighted in the REF.
Rater 4 Rating = 4
IR Justification: The AI version highlights insightful aspects, such as children's agency in crisis management and predicts potential future policy shifts. However, the original REF document provides specific examples of how children's voices have led to concrete policy changes, which adds depth to the analysis that the AI-generated impacts do not fully capture.

"
"56. Study Title : Improved Crab Fisheries Management Benefits Coastal Livelihoods in Brazil
Rater 1 Rating = 4 - Very Good
IR Justification: The REF submission provides excellent insight into the direct path to policy change and the detailed technical aspects of the forecasting tool for crab mass mating. The AI-generated impacts add value by discussing potential future directions for global applications of the model and cooperative structures, which demonstrates forward-thinking analysis. Both versions are insightful, though the REF submission is more focused on current outcomes while the AI-generated version offers additional predictions for long-term scalability.
Rater 2 Rating = 4 (Very Good)
IR Justification: The AI-generated content is insightful, particularly in its projection of long-term sustainability impacts and the broader application of the fisheries management tool. It also suggests future trends, such as how digital tools like the REMAR_CIDADÃO app might evolve. However, these insights do not introduce entirely new concepts that build extensively on the REF content.
Rater 3 Rating = 4 - Very Good
IR Justification: The AI-generated impacts offer valuable insights into the potential transferability and scalability of the fisheries management model, along with future implications for economic stability and environmental sustainability. The REF version is insightful but remains more specific to immediate impacts and policy outcomes in Brazil.
Rater 4 Rating = 4

IR Justification: The AI-generated report introduces potential extensions of the current impact to future trends, such as using similar predictive tools in other biodiversity-related fishery management scenarios. This adds an insightful dimension beyond the immediate context. However, the REF version includes a more profound narrative of stakeholder collaboration and direct quotes, which provide contextually rich insights into community and governmental interactions.

"
"57. Study Title : Influencing Organisational Strategy to Support Responsible Business Practice
Rater 1 Rating = 3 - Good
IR Justification: The AI-generated content provides valuable insights, particularly around the long-term policy implications of the research, but it tends to generalize future trends without delving deeply into specific potential outcomes or stakeholder responses. The REF version includes direct quotes and feedback from key participants, adding depth to understanding the research's societal and practical impacts.
Rater 2 Rating = 3 (Good)
IR Justification: The AI-generated report demonstrated solid insightfulness, especially in suggesting potential future impacts and scalability of the research. It presented valuable implications for extending OR methodologies to diverse sustainability challenges. However, it did not provide significant new or deeper insights into the specific societal outcomes or stakeholder behaviors detailed in the REF submission, where the impacts were grounded more in established relationships and specific events.
Rater 3 Rating = 5 (Excellent)
IR Justification: The AI-generated impacts added depth by suggesting how the research could influence future sustainable business practices and governance frameworks, showing proactive and strategic foresight.
Rater 4 Rating = 4: Very Good
IR Justification: The AI content provided valuable insights into the future potential of these programs, suggesting that responsible business practices may influence future corporate strategies and public policy, which adds depth beyond the immediate impact outlined in the REF version.

"
"58. Study Title : Billmonitor: predicting the best mobile phone contract for businesses and individual users
Rater 1 Rating = 4: Very Good
IR Justification: The AI-generated report offers insightful predictions about potential future applications of the research, particularly in the context of technological transferability and innovation. It correctly identifies the disruptive potential of Billmonitor in transforming mobile contract pricing and predicts further integration into financial software. While the REF submission is more focused on concrete, realized impacts, the AI version provides a forward-looking analysis of broader academic and technological applications.
Rater 2 Rating = 5 - Excellent
IR Justification: The AI-generated content provides highly insightful predictions on future trends, such as how the Billmonitor algorithm could apply to other fields (e.g., financial market modeling and personalized healthcare in genomics). The AI-generated impacts extend beyond the original REF submission by offering future-oriented insights, theoretical expansions, and suggestions for cross-disciplinary applications. These predictions are forward-thinking and insightful, adding significant value by expanding the potential impact of the research.
Rater 3 Rating = 3 - Good
IR Justification: ChatGPT provides forward-looking implications, including projections of user behavioral impacts on telecom cost structuring and the potential for algorithmic adaptations in other utilities. However, the REF submission provides more tangible insights through direct evidence of cost savings and business adoption statistics, adding concrete measures of impact.
Rater 4 Rating = 3 (Good)
IR Justification: While the AI-generated report is insightful in predicting broader implications of algorithmic models, it falls short in adding actionable insights on operational refinements that could be considered for Billmonitor. The REF document, however, includes specific user testimonials and direct applications that enhance the understanding of real-world usability.

"
"59. Study Title : Mediating Modern German: reaching new and diverse audiences through translation, engagement, and performance
Rater 1 Rating = 4 - Very Good
IR Justification: The ChatGPT-generated impact provides insightful reflections on how the research might influence interdisciplinary collaboration and future studies in Germanic poetry and translation studies. It expands on potential technological impacts, such as digital

archiving, that are not explicitly mentioned in the REF version, providing forward-thinking considerations.
Rater 2 Rating = 4 (Very Good)
IR Justification: The AI-generated impacts are insightful, particularly in their ability to extrapolate potential future trends and implications of the research. It mentions the possibility of further public policy engagement and technological innovations in poetry dissemination, which extend beyond what is explicitly covered in the REF submission. This demonstrates a good understanding of the broader implications of the research. However, the AI version does not dive as deeply into some of the nuanced cultural impacts as the REF version does.
Rater 3 Rating = 3 (Good)
IR Justification: The AI-generated impacts provide a good overview and are insightful in framing the importance of cultural translation and poetry accessibility in modern society. However, it does not fully extrapolate future trends or implications as deeply as the original submission, which provides more robust reflections on the long-term influence of German poetry in English-speaking contexts.
Rater 4 Rating = 3
IR Justification: The AI-generated impacts provide a meaningful summary but lack some actionable insights present in the REF submission, such as identifying specific audience engagement tactics that could be used to deepen impact in underrepresented communities. The AI model does suggest further pathways, which may support future project planning.

"
"60. Study Title : OMass Therapeutics: New technology for drug discovery with economic benefit to the UK
Rater 1 Rating = 5 (Excellent)
IR Justification: The AI-generated content provides insightful predictions about the future applications of the research, particularly in terms of how mass spectrometry techniques might advance pharmaceutical research and healthcare policies. It also extends the discussion to the potential global impacts on cancer treatment, which demonstrates a high level of foresight.
Rater 2 Rating = 3: Good
IR Justification: The AI-generated content showed good insight into the future applications and interdisciplinary potential of the research but lacked some of the deeper foresight regarding the strategic industry partnerships and venture capital growth that the REF submission covered in depth.
Rater 3 Rating = 5 - Excellent
IR Justification: The AI-generated version provides insightful interpretations by projecting potential future impacts, including enhancing international healthcare strategies and advancing diagnostic methods. This prediction goes beyond the REF submission, highlighting broader long-term impact trajectories in public health and biophysics.
Rater 4 Rating = 3 - Good
IR Justification: The AI-generated impacts were relevant and offered an adequate interpretation of potential implications. Although the AI did not predict direct future industry trends or propose novel directions based on the data, it did extrapolate well on the core benefits of the mass spectrometry method. Further foresight on industry or clinical adoption would have enhanced the insightfulness rating.

"
"61. Study Title : Living With Feeling: Transforming Understandings of Emotional Health
Rater 1 Rating = 3
IR Justification: While the AI-generated version provides relevant and accurate content, it tends to stay within the bounds of the information available in the REF submission without significantly expanding beyond it. The original REF version offered some reflection on the wider social and policy implications, while ChatGPT could provide more forward-looking insights or innovative ideas for future application.
Rater 2 Rating = 5 (Excellent)
IR Justification: The AI-generated impacts demonstrated strong insightfulness, especially by predicting future societal and educational trends based on the research's findings. The analysis of how emotional vocabularies could influence future public health, education, and social policies added a forward-thinking dimension that extended beyond the immediate impacts covered in the REF submission.
Rater 3 Rating = 3
IR Justification: While insightful, the AI-generated content tends to generalize outcomes, such as policy impacts, without deeply exploring specific applications or societal shifts as detailed in the REF submission. Further predictions about future influence in healthcare and social reform could add depth.

Rater 4 Rating = 4 (Very Good)
IR Justification: The AI-generated impacts offer valuable insights, especially in predicting the scalability of the emotional health curriculum and suggesting broader applications in healthcare and policy settings. While the REF submission is richly detailed, the AI analysis provides additional foresight, particularly regarding long-term societal and educational transformations. However, it lacks specific historical context and reflections on cultural implications, such as how emotional vocabularies have historically shaped public attitudes—a detail emphasized in the REF.

"
"62. Study Title : Litigation as a Tool to Support Social Change: Indigenous Peoples, Human Rights, and Legal Empowerment
Rater 1 Rating = 4: Very Good
IR Justification: The AI version provides insightful analyses, especially regarding the potential scalability of litigation strategies and how they might influence future legal reforms globally. The AI-generated content predicts long-term global impacts on indigenous rights, which extend beyond the immediate case studies discussed in the REF submission.
Rater 2 Rating = 3 (Good)
IR Justification: The AI-generated impacts are insightful, providing meaningful analysis of how the research could influence future legal strategies. However, it lacks some of the nuanced insights present in the REF submission regarding the community-level empowerment effects and the specific outcomes from stakeholder engagements.
Rater 3 Rating = 4: Very Good
IR Justification: The AI version provides valuable insights by anticipating long-term implications of litigation on future policy frameworks, sustainability, and ethical standards. It effectively builds on the REF findings with forward-looking insights, although it lacks some specific case study-based insights about litigation's immediate benefits to indigenous communities.
Rater 4 Rating = 4: Very Good
IR Justification: ChatGPT provided future-focused implications by predicting broader impacts of indigenous litigation as a potential global precedent. The REF submission focused more on current, tangible outcomes but less on future applications outside the documented communities, demonstrating a balance of insight between present and predictive relevance.

"
"63. Study Title : Accelerating the development of medicines for children through an open-access excipient database
Rater 1 Rating = 3 - Good
IR Justification: ChatGPT offers insights into how the STEP database can affect future research and regulatory frameworks, especially in its mention of international collaborations and potential policy changes. However, it does not predict future economic or regulatory trends as explicitly as the REF-submitted version.
Rater 2 Rating = 4 (Very Good)
IR Justification: The AI-generated impacts provide valuable insights, especially regarding the potential for the STEP database to influence future drug development processes and regulatory trends. These insights go beyond the explicit outcomes described in the REF submission and offer a more forward-looking perspective.
Rater 3 Rating = 4 (Very Good)
IR Justification: The ChatGPT-generated impact offers insightful extensions, especially regarding the broader public health and ethical considerations of making pediatric formulations safer, predicting STEP's role in driving pediatric research standards. It generally aligns with the REF submission, though it misses certain nuanced applications, like the specific excipient reduction practices cited by Proveca.
Rater 4 Rating = 3 (Good)
IR Justification: While insightful, the ChatGPT version primarily described the impacts without deeper implications or trend analysis for future pediatric pharmaceutical innovations. For example, the REF submission noted the direct cost and time savings that STEP provided, but the ChatGPT version lacked predictive insights about future savings or technological integration with other databases.

"
"64. Study Title : Pluralistic Evidence for Successful Policymaking about Reactive Systems
Rater 1 Rating = 3 (Good)
IR Justification: The ChatGPT-generated impacts provide a meaningful projection of future trends, particularly in scalability and transferability of IQ frameworks. However, it falls short in extending actionable insights specific to the ethical dimensions that were central to

the REF submission, such as data privacy and public confidence in AI systems.
Rater 2 Rating = 3 - Good
IR Justification: While the AI-generated content accurately represents many aspects of the study, it lacks the nuanced foresight and deep understanding of the long-term implications of the research that the REF submission offers. The original submission provides a deeper reflection on how evidence plurality and mechanism assessments have already transformed policy frameworks and how they continue to influence policy beyond publication. The AI-generated version, while robust, provides less detail on these evolving dynamics and offers fewer unique predictions or inferences about future impacts.
Rater 3 Rating = 3 (Good)
IR Justification: ChatGPT's content provides insightful connections between ethical AI frameworks and health policy, extending potential applications. However, it lacks forward-looking implications related to legislative changes or future NICE/IARC methodological shifts noted in the REF submission.
Rater 4 Rating = 3: Good
IR Justification: The AI version provides insightful interpretations of the research's future applications, particularly in ethical frameworks and emerging AI policy needs. While it forecasts potential impacts in data governance and ethics, it does not fully predict the nuanced and varied implications that arose from individual stakeholder actions as precisely as the REF submission, such as the nuances in NHS frameworks discussed in the REF.

"
"65. Study Title : Reshaping professional heritage practice and changing understanding of heritage in the UK and internationally
Rater 1 Rating = 3 (Good)
IR Justification: While the AI-generated content is insightful, offering future-oriented discussions and broad applications in heritage management, it lacks some of the in-depth professional and stakeholder feedback found in the REF submission. The AI version highlights emerging trends but misses testimonials that offer richer, practical insight.
Rater 2 Rating = 3 (Good)
IR Justification: The AI-generated report demonstrates a reasonable level of insight, predicting potential future impacts on international heritage policy and public behavior changes regarding heritage appreciation. However, it lacks depth in exploring the ongoing challenges and potential areas of future collaboration that were identified in the REF submission.
Rater 3 Rating = 3 - Good
IR Justification: The AI-generated content provides insightful observations on the potential cross-domain impacts and sustainable policy applications of the heritage research, suggesting future trends like integrating cultural and natural heritage in climate adaptation strategies. However, some reflections found in the REF submission about organizational changes (e.g., shifts in perception among museum staff) provide nuanced insights not covered by the AI-generated text.
Rater 4 Rating = 3 (Good)
IR Justification: The AI-generated content provides meaningful insights into the societal role of heritage preservation and identifies potential cross-disciplinary influences in conservation. It predicts broader implications for cultural policy and engagement, offering a macro perspective that extends beyond immediate project impacts. However, the REF case presents a more nuanced exploration of each engagement, such as specific shifts in institutional perspectives and the impact of artist collaborations, which offer deeper contextual insights.

"
"66. Study Title : Shaping the legal framework for Brexit
Rater 1 Rating = 4 (Very Good)
IR Justification: ChatGPT-generated impacts demonstrate insight into how the research might influence future legal discourse, particularly by extending beyond Brexit to broader constitutional and public law debates. However, it doesn't predict long-term legal transformations as deeply as the REF version.
Rater 2 Rating = 4 (Very Good)
IR Justification: The AI-generated content shows a high level of insight, especially in predicting broader implications of the legal frameworks shaped during Brexit. It emphasizes future trends, such as the potential application of these legal principles in other EU-related constitutional debates. However, the depth of contextual analysis present in the REF submission is somewhat lacking in the AI version.
Rater 3 Rating = 4 (Very Good)
IR Justification: The AI-generated report provides meaningful insights that support a deeper understanding of the legal impacts, especially in terms of policy influence and societal awareness. However, while it makes connections to broader implications, it lacks some nuanced historical context provided in the REF submission, which adds further depth to the legal

arguments.
Rater 4 Rating = 4 (Very Good)
IR Justification: The AI-generated content displayed considerable insight by extrapolating the research's potential impact on future legal debates regarding parliamentary authority. It identified potential trends in UK constitutional law that could evolve from the Brexit judicial rulings, demonstrating an understanding of the broader implications of the cases. However, while insightful, it occasionally lacked specific examples illustrating these extrapolations, which would provide more immediate relevance.

"

"67. Study Title : Bristol's materials research is keeping the UK's fleet of nuclear power stations safe and operating.
Rater 1 Rating = 4
IR Justification: The AI-generated content provides insightful extensions of the impact beyond what is explicitly stated in the REF submission. It forecasts potential future applications of the research in other high-temperature, high-stress environments and hints at future developments in nuclear technology beyond the scope of the specific reactors discussed in the REF submission.
Rater 2 Rating = 3 (Good)
IR Justification: While the ChatGPT-generated version shows understanding of the impact's immediate academic and economic effects, it lacks deeper predictive insights into future trends in nuclear safety or potential policy shifts due to UoB's research. The REF submission provided more concrete examples of immediate results, such as the Hunterston B Reactor safety extension.
Rater 3 Rating = 4: Very Good
IR Justification: The AI-generated content is insightful, providing a forward-looking perspective on how UoB's findings could shape future safety standards and technological advancements. It offers additional insights on transferring findings to other industries and potential adaptations for new reactor technologies, though some contextual specificities in the REF submission provide additional value.
Rater 4 Rating = 4 (Very Good)
IR Justification: The ChatGPT version demonstrates insight into the implications of material degradation in high-stress environments, suggesting future applications and scalable benefits for broader industrial use. It emphasizes predictive models that could influence future standards in nuclear and other high-stress industries.

"

"68. Study Title : Radical advance in treating age-related macular degeneration leading to global impact in prevention of blindness
Rater 1 Rating = 4 (Very Good)
IR Justification: The ChatGPT-generated research impacts demonstrate insightfulness by predicting future trends and suggesting implications that extend beyond the immediate research findings, such as its potential influence on the pharmaceutical industry and its application in low-income countries. However, the REF submission provides more direct, specific examples of immediate impacts, such as its role in shaping NICE guidelines and influencing judicial decisions in the UK.
Rater 2 Rating = 4 - Very Good
IR Justification: The AI-generated content displayed strong insight by addressing potential future applications and scalability of the findings, particularly the emphasis on cost-effectiveness and its applicability to healthcare systems in low-resource settings. It extended the analysis of impact beyond what was explicitly stated in the REF submission, demonstrating forward-looking projections.
Rater 3 Rating = 3 (Good)
IR Justification: The AI content provides useful insights, predicting long-term global applications and broad policy implications of cost-effectiveness for AMD treatments. However, it lacks certain in-depth projections present in the REF submission, such as specific future roles of policy adoption in Europe and the cost implications for multi-level healthcare systems.
Rater 4 Rating = 4 (Very Good)
IR Justification: The AI version presents insightful generalizations that could help predict the global uptake of cost-effective AMD treatment practices. While not as regionally detailed as the REF, the AI summary considers pathways for future scaling and addresses sustainability, which may facilitate economic policy adaptation.

"

"69. Study Title : Rolling programme of research, centred on the National Joint Registry, to improve the outcomes of hip and knee replacements worldwide
Rater 1 Rating = 4 - Very Good

IR Justification: The AI-generated content provides some insightful projections about the future impact of the study, particularly regarding the long-term changes in prosthetic technology, interdisciplinary effects on materials science, and future healthcare policies. It is somewhat predictive, which adds to its insightfulness. However, the REF version provides more detail on the real-world outcomes that have already occurred, such as NHS cost savings and the reduction in perioperative mortality rates. The REF submission is more grounded in currently observed impacts.
Rater 2 Rating = 3 (Good)
IR Justification: The AI version provides reasonable insights into future impacts, particularly in its projections for long-term implant safety and regulatory updates. However, the REF submission offers more immediate actionable insights, such as the financial and clinical implications of implant recommendations, which the AI version does not fully capture.
Rater 3 Rating = 3
IR Justification: The AI version presents several insightful projections, such as the broader environmental impact of safer implants, but it does not predict patient care outcomes or future research directions with as much specificity as the REF. The REF team's insights into cost implications and clinical outcomes over a decade provide a sharper forecast for direct impacts.
Rater 4 Rating = 4
IR Justification: The ChatGPT version anticipates future trends, including shifts in implant technology and patient care standards. It offers potential directions for technology advancement and ethical considerations, reflecting the broader implications of the UoB research well. Some specific policy and technical insights from REF are less detailed.

"
"70. Study Title : Reducing breast and ovarian cancer occurrences in women at high risk
Rater 1 Rating = 4 (Very Good)
IR Justification: The AI-generated content demonstrates insight by discussing future trends, such as expanding genetic screening access and integrating new genetic loci into breast cancer risk prediction models. This aligns well with REF's emphasis on ongoing improvements to BOADICEA. However, the AI version could be enhanced by anticipating specific international policy adoptions, as observed in the REF version's impact on various national healthcare systems.
Rater 2 Rating = 3 - Good
IR Justification: The AI content is adequately insightful in exploring the BOADICEA tool's broader implications, including enhanced patient autonomy in health decisions and hypothetical future applications. However, it lacks some of the evidence-backed, specific outcomes seen in the REF submission, which cites mortality reduction and preventive intervention statistics tied to BOADICEA's real-world use. Insightfulness could be improved by including similar quantified impacts or by forecasting specific future applications based on current trends.
Rater 3 Rating = 4
IR Justification: The ChatGPT impacts show insightfulness by predicting BOADICEA's role in future personalized medicine, broader genetic counseling applications, and policy shifts toward preventive health models. However, while innovative in its policy and clinical applications, it lacks the level of granularity seen in the REF submission regarding risk stratification for specific populations.
Rater 4 Rating = 3 (Good)
IR Justification: ChatGPT provided insightful extrapolations on how BOADICEA could potentially shape future screening policies and improve healthcare efficiencies. However, it lacked predictive depth regarding the evolving landscape of genetic testing and implications on regulatory standards, which were more implicitly addressed in the REF case.

"
"71. Study Title : The Haydn Scale: Changing policy and practice for improving pupil behaviour in schools
Rater 1 Rating = 3 (Good)
IR Justification: While insightful in presenting the broader implications of Haydn's research, the AI-generated content does not consistently capture the in-depth perspective of societal and behavioral changes reflected in the REF document. Notably, the AI version includes foresight into tech integration in classroom management but lacks a deep exploration of institutional stakeholder influence.
Rater 2 Rating = 3 (Good)
IR Justification: The AI version brings some insightful suggestions for future research directions, such as potential long-term impacts on international policies. While it aligns with REF's long-term goals, it does not provide as much depth on institutional training impacts or policy engagement, which could strengthen the depth of insights on actual influence on educational policy shifts.
Rater 3 Rating = 4 (Very Good)

IR Justification: ChatGPT provided insightful observations regarding the potential for broad-scale integration of the Haydn Scale and inferred benefits for teacher resilience in classroom management. The AI insights, however, could further elaborate on systemic educational policy changes and the future of teacher training integration.
Rater 4 Rating = 3 (Good)
IR Justification: The ChatGPT version demonstrates moderate insightfulness by extending the impact discussion to future implications, such as the long-term potential of improved classroom climates. It also draws connections between the Haydn Scale and teacher well-being indirectly. However, the original REF submission is more detailed in presenting unique insights, like testimonials from specific trainee experiences and detailed stakeholder perspectives, making the REF insights richer in context.

"
"72. Study Title : ""Definitive demonstration of the adverse cardiovascular effects of air pollution leads to policy change at the local, national, and international level""
Rater 1 Rating = 4 - Very Good
IR Justification: The AI-generated version demonstrates insight by identifying potential future areas of research impact and community engagement, such as sustained education initiatives and public health collaborations. However, while it forecasts long-term adoption pathways, it does not specify certain high-impact events and timelines discussed in the REF submission.
Rater 2 Rating = 3 - Good
IR Justification: The AI version offers forward-looking predictions on air quality improvements' long-term health impacts, contributing to policy and public health insights. While the REF focuses on concrete outcomes and realized impacts, such as new legislative priorities, the AI version adds speculative insights that are forward-thinking. However, it lacks immediate contextually rich insights found in the REF case studies, such as current media or governmental alignment.
Rater 3 Rating = 4 (Very Good)
IR Justification: The AI-generated impacts exhibit foresight by suggesting broader implications of the research for future environmental policies and health interventions, even predicting potential applications in global public health frameworks. However, this insight is more hypothetical and less grounded in specific examples than the REF submission, which provides actual impacts achieved.
Rater 4 Rating = 5: Excellent
IR Justification: The AI version provided unique insights by anticipating the scalability of the findings to influence global policy trends and offered potential future implications, such as advocating for stricter global air quality standards based on health data extrapolation. This forward-looking view presents a robust interpretation of long-term research significance.

"
"73. Study Title : Navigating Inclusion in International Peace Processes
Rater 1 Rating = 4 (Very Good)
IR Justification: The ChatGPT content demonstrates insightful analyses, particularly in terms of the potential long-term impacts of the research, such as adapting the framework for different peace process settings globally. It highlights unique aspects of technological adaptation and scalability, anticipating future research and policy development needs, which go beyond the direct impact statements in the REF submission. However, it misses some of the specific real-world impact projections such as expected outcomes for local communities involved in conflict-affected areas, which are highlighted in the REF document.
Rater 2 Rating = 3 – Good
IR Justification: The ChatGPT-generated report provides meaningful insights into the broader implications of the PSRP's contributions to peace processes, particularly regarding the cross-disciplinary and global relevance of the PA-X database. However, the REF version offers additional nuanced insight, including specific long-term applications in policy and feedback from UN stakeholders that project future policy shifts. The ChatGPT version identifies potential trends, though its insights are not as detailed in certain application areas, such as direct peacebuilding actions or procedural stages in policy implementation, limiting its scope of insight.
Rater 3 Rating = 3 (Good)
IR Justification: ChatGPT-generated impacts provide some valuable predictive insights, particularly in terms of expanding the PA-X Database's application and its potential relevance in future policy frameworks. However, it lacks the nuanced foresight found in the REF submission concerning incremental policy impacts (e.g., progressive gender inclusivity steps over time) and specific recommendations for conflict resolution frameworks in fragile states. This limits the insightfulness to a "Good" rating as it adequately presents insights but lacks the depth of practical, forward-looking scenarios found in the REF.

Rater 4 Rating = 3 (Good)
IR Justification: The AI report brings forward some forward-looking insights on the influence of gender policies in peace processes globally, highlighting potential pathways and scalability in diverse contexts. However, these insights are relatively broad compared to the detailed predictive aspects found in the REF 2021, which discusses specific downstream effects on peace negotiations and future UNSC mandates for gender inclusion.

"
"74. Study Title : Documenting and Protecting Survivors of Torture and Ill-Treatment Living in Poor Communities
Rater 1 Rating = 4 (Very Good)
IR Justification: The REF 2021 version is insightful in its operational insights, showcasing real-world applications that directly addressed gaps in human rights documentation. This includes specific data-driven examples, such as the use and impact of mobile apps in Kenya. ChatGPT's insights extend further by suggesting potential cross-disciplinary adaptations and broader applications in digital and policy fields, which are insightful but less immediately actionable.
Rater 2 Rating = 4 (Very Good)
IR Justification: The AI-generated impacts demonstrate good insight by anticipating potential applications of the findings, such as applicability in other low-income regions and broad policy recommendations for similar organizations. Additionally, the AI version identifies how documenting practices could evolve, possibly influencing future human rights tools and methodologies. This added depth aligns well with the REF impacts, although some specificity is lacking.
Rater 3 Rating = 4 (Very Good)
IR Justification: The AI-generated version demonstrates insight by suggesting potential long-term implications and adaptations for human rights practices. It recognizes the need for systemic reform in human rights documentation and mentions the scalability of practices established in the research.
Rater 4 Rating = 5
IR Justification: The AI-generated version provided insightful forecasts, including potential uses of emerging technologies to enhance documentation accessibility, and offered nuanced implications for future documentation practices. This predictive value extended the REF submission's findings by suggesting long-term adaptations and growth areas.

"
"75. Study Title : Transforming genomic selection in commercial breeding programmes for pigs, dairy goats, and poultry
Rater 1 Rating = 4 (Very Good)
IR Justification: The AI impacts demonstrate insight into future trends and implications, emphasizing genomic technology scalability and cross-species applications that align with broader genomic and agricultural advancements. The REF version, however, offers more immediate, grounded impacts specific to the livestock industry, such as direct breeding program improvements and economic returns. While the AI content introduces future-oriented implications, adding similar practical insights would deepen the overall impact.
Rater 2 Rating = Rating: 4 (Very Good)
IR Justification: The AI-generated report provides valuable insights into the potential long-term implications of genomic advancements for future agricultural practices and sustainability, which enhances its value beyond merely descriptive content. The AI version projects broader societal and food security benefits that could result from continued genomic selection enhancements, adding a future-oriented layer to the analysis. This foresight makes the AI-generated content highly insightful and forward-thinking, although specific references to current, short-term impacts, such as productivity increases achieved in collaboration with Genus, Aviagen, and YDG, are less prominently discussed. Including these metrics would add a tangible dimension to the insightfulness of the AI-generated content.
Rater 3 Rating = 4: Very Good
IR Justification: AI content captures essential long-term effects and potential expansion of genomic selection's impact. Insightfully, it hints at sustainability and efficiency benefits, aligning these with the overarching theme of resource optimization. However, AI insights are less detailed in predicting economic metrics and specific advancements in genetic marker accuracy.
Rater 4 Rating = 3 (Good)
IR Justification: The AI-generated impacts provide a thoughtful perspective on economic and technological advancement but fall short on specific long-term forecasting or policy implications that the REF document touches upon. Insight on potential applications beyond the study's immediate scope is present but lacks the specificity found in REF submission's targeted

predictions.

"
"76. Study Title : Employing polymer physics to improve gluten-free bread structure
Rater 1 Rating = 3 (Good)
IR Justification: The AI report gives insightful connections to additional technological
advancements and possible future applications in the polymer industry, showing moderate
foresight. However, it lacks some of the specific innovation pathways, such as the cross-linking
density adjustments and subtractive formulation approaches emphasized in the REF, which reflect
the project's foundational knowledge in ingredient stability.
Rater 2 Rating = 4 (Very Good)
IR Justification: The AI-generated report offers insightful predictions, particularly regarding
potential future economic impacts and scalable technology applications, suggesting broader
industrial relevance. The report demonstrates an understanding of the transformative effects on
gluten-free product development beyond direct economic benefits. However, it does not delve into
predictive insights on academic or entrepreneurial frameworks to the extent the REF submission
achieves with mentorship outcomes and networking initiatives.
Rater 3 Rating = 4
IR Justification: The AI-generated content provides insightful future directions and
implications for the GF bread sector, anticipating trends in recipe stability and market
potential. While strong, certain real-time practicalities and consumer feedback mechanisms were
less prominent compared to REF.
Rater 4 Rating = 4 (Very Good)
IR Justification: The AI-generated impacts provide insightful future possibilities for using
polymer physics in food technology, especially in predictive applications for ingredient
performance. However, the REF submission presents more actionable insights, such as direct
commercial applications and institutional entrepreneurship, offering concrete examples of impact
rather than broader, speculative insights.

"
"77. Study Title : Peripheral Impressionisms: challenging perceptions of Impressionism
Rater 1 Rating = 4
IR Justification: The AI-generated impacts are insightful in foreseeing how the exhibitions
might influence future art market trends and continue impacting public interest in lesser-known
Impressionist artists. This interpretation extends the core objectives stated in the
REF-submitted impacts, projecting longer-term cultural relevance. However, the REF submission
offers deeper insights into specific curatorial strategies, making it slightly more instructive
in detailing practical, immediate effects within art institutions.
Rater 2 Rating = 4
IR Justification: The AI-generated impacts offer insightful projections regarding the potential
future influence of peripheral impressionism on global curatorial approaches, capturing the
project's influence on both public perception and art market trends. This foresight is valuable,
as it provides potential pathways for future impact, although it does lack specific predictive
metrics discussed in the REF case study.
Rater 3 Rating = 3 (Good)
IR Justification: The AI version offers some insight into the broader cultural and societal
impacts of the exhibitions but falls short of providing predictions on future curatorial
practices or shifts in academic discourse, which are implied but not expanded upon. Insight on
Daubigny's impact could further contextualize cultural shifts in Impressionism's reception.
Rater 4 Rating = 3 (Good)
IR Justification: AI interpretations suggest insightful future projections, particularly on
sustained cultural relevance and economic potential for peripheral artists' work. However, it
does not fully address the unique influence on exhibition practices that REF's details on
changes in institutional approaches underscore.

"
"78. Study Title : Raising the profile of Scottish Literature through writing and consultancy
Rater 1 Rating = 4 - Very Good
IR Justification: The AI-generated impacts provide insightful extensions of the original
impacts, particularly by envisioning how similar cultural projects could be implemented globally
and predicting potential social and economic outcomes from promoting regional literature
internationally. While insightful, some predictions lack clear grounding in the original study
context.
Rater 2 Rating = 3
IR Justification: The AI-generated content provides some insightful connections, such as
potential educational applications of the project. However, it lacks the nuanced interpretation

of the REF submission regarding cultural implications and historical context in promoting Scottish literature. For instance, the AI version mentions the general impact of the RBS project on Scottish cultural awareness but misses the REF submission's insight on the generational significance of such projects for future Scottish identity and the nuanced use of historical figures as symbols of cultural pride.

Rater 3 Rating = 4 (Very Good)

IR Justification: The AI version brings additional insight into potential long-term benefits of raising awareness in literature and cultural tourism beyond Scotland, which could encourage broader global interest in Scottish literature. This demonstrates future relevance and possible adaptability to other contexts. However, it does not include all of the nuanced, direct feedback on impacts (like those from the Nile Design or gallery directors), limiting insight into specific stakeholder responses.

Rater 4 Rating = 3

IR Justification: The AI-generated impacts provided meaningful insights, emphasizing the project's potential for extended influence in educational settings and further cultural engagement. While insightful, these projections lacked the depth of future-oriented cultural policy implications that might stem from deeper literary integration, which the original submission implied through sustained literary consultancy. The AI version did not predict specific cultural evolution trends that could emerge from public exposure to Scottish literature as envisioned by the research team.

"

"79. Study Title : Secukinumab becomes the first interleukin-17A inhibitor approved for psoriatic arthritis

Rater 1 Rating = 5 (Excellent)

IR Justification: The ChatGPT-generated version offers highly insightful interpretations of the research impact by projecting broader future applications and suggesting additional potential stakeholders, including underserved patient populations and emerging markets. It suggests that the research's impact could evolve over time, incorporating aspects such as increased accessibility and ethical considerations.

Rater 2 Rating = 4 (Very Good)

IR Justification: The AI-generated content displays significant insight, particularly in projecting the broader implications of IL-17 inhibition beyond PsA treatment, which adds a layer of prospective value not fully developed in the REF submission. However, while these projections are valuable, they lack the empirical specificity and clinical trial outcomes present in the original. The REF submission's insight into future directions is embedded more in empirical study extensions and patient-reported outcomes, which gives it a data-driven predictive quality that enhances insight.

Rater 3 Rating = 3 - Good

IR Justification: The AI-generated impact report includes insightful comments about secukinumab's potential application to other inflammatory diseases and briefly suggests potential future pathways in autoimmune therapy. However, it lacks predictive insights on policy changes or cost-saving measures that are found in the REF version.

Rater 4 Rating = 3

IR Justification: The AI-generated impact provides a solid synthesis but lacks depth in foresight into future trends beyond PsA and misses nuanced discussions around potential market expansion or the future direction of biologic therapies. In contrast, REF highlights insights into the expanding role of Secukinumab in various clinical recommendations, underscoring a deeper understanding of potential healthcare implications.

"

"80. Study Title : Establishing the Facts, Developing Professionalisation and Enabling Transparent 'Pilgrim-centred' Communication in the UK Hajj Sector

Rater 1 Rating = 5 - Excellent

IR Justification: The AI-generated impacts demonstrated insightfulness by linking UK Hajj governance issues to global challenges in religious tourism regulation and professionalization. The AI version suggested possible future uses, such as governance frameworks applicable to broader religious tourism sectors, anticipating potential adaptations of findings across cultural sectors. The AI impact summary recognized that findings could inform policy beyond the immediate UK context, reflecting insight into broader applications.

Rater 2 Rating = 4 (Very Good)

IR Justification: The ChatGPT-generated report demonstrates strong insights, capturing the project's relevance and forecasting the implications of McLoughlin's work on policy and commercial practices in the Hajj industry. It adds predictive analysis on the UK Hajj sector's future needs and emerging trends, showing a deep understanding of the research's potential impacts. However, the original REF report provided more in-depth examples and personal

stakeholder testimonies that contextualize the impacts on individual pilgrims and organizers.
Rater 3 Rating = 4 (Very Good)
IR Justification: The ChatGPT version provides predictive insights, such as the potential for Hajj governance frameworks to be applied to other religious tourism contexts. It highlights the scalability of transparency and governance mechanisms beyond the Hajj industry, showing awareness of broader industry implications. This is in line with the REF submission's impact on the sector's transparency but does not cover the nuanced governmental and legislative context detailed in the REF submission.
Rater 4 Rating = 4 - Very Good
IR Justification: The AI-generated impacts provide insightful connections, particularly in highlighting the ethical challenges within the industry, such as transparency issues and consumer protection, which are in line with McLoughlin's findings. The AI version could, however, add further predictive insights regarding future regulatory frameworks in the Hajj sector. It effectively identifies the ongoing need for professionalization, offering a structured perspective on sector challenges.

"

"81. Study Title : A Bridge for Spies: Overcoming the Practitioner-Academic Gap in Intelligence and Security
Rater 1 Rating = 3 - Good
IR Justification: While the AI report offers valuable insight into future applications, especially the potential for adapting methods across nations, it does not delve deeply into the foundational impact or long-term policy shifts as in the REF. For example, it did not consider the potential challenges in implementation or the nuanced changes in agency protocols due to Dover's work. The AI version does, however, add value by projecting future uses of the research in broader security frameworks.
Rater 2 Rating = 3: Good
IR Justification: While the AI-generated impacts provide reasonable foresight into future implications of academic-intelligence collaboration, they lack some of the deeper insights seen in the REF submission, such as the potential cascading effects of professionalized intelligence analysis on international intelligence standards and sustained security improvements. However, the AI version's insights on potential societal benefits (e.g., public trust and ethical intelligence practices) extend the REF submission's outcomes.
Rater 3 Rating = 4
IR Justification: The ChatGPT version provides forward-looking insights, especially regarding how the research might inform ongoing cross-sectoral collaboration beyond intelligence. It offers valuable context for the sustainability of academic engagement in government intelligence, which complements the foundational impacts emphasized in the REF submission.
Rater 4 Rating = 4: Very Good
IR Justification: The AI-generated impacts offer a forward-thinking view by emphasizing the potential for scalability across policy sectors and anticipating trends such as interdisciplinary applications of the framework in other policy areas. However, the REF version offers testimonials and agency feedback, which add grounded insight into the practical results of the engagement. While the ChatGPT version is imaginative and relevant, the REF version includes actionable insights, making it highly relevant for current policy implications.

"

"82. Study Title : Influencing Labour Standards and Stakeholder Action Through International, European and National Law and Policy
Rater 1 Rating = 4 (Very Good)
IR Justification: The ChatGPT-generated report demonstrates insightfulness by highlighting broader applications of the research findings beyond direct policy influence, such as predicting the possible adoption of similar labor reforms in other EU member states and drawing connections to social stability and legal frameworks. Additionally, it suggests further applications of the research for addressing worker rights and inclusive labor standards in varied economic climates. The AI-generated version, however, is slightly less detailed in suggesting specific future trends and organizational impacts (e.g., specific policy adoptions).
Rater 2 Rating = 3: Good
IR Justification: The AI-generated impacts demonstrate a good level of insight, discussing potential long-term implications of the study, such as labor policy shifts in EU nations experiencing financial crises and the need for coordinated social dialogue frameworks. However, it lacks some predictive insights seen in the REF version, such as specific legal and economic forecasts for impacted countries.
Rater 3 Rating = 4
IR Justification: The AI-generated impact provides meaningful insights into future applications and the broader implications of labor policy evolution, though slightly less detailed in the

unique historical and political context emphasized in the REF submission. The AI's predictive insights regarding long-term policy implications align closely but could include more explicit connections to subsequent EU labor standards.
Rater 4 Rating = 3 (Good)
IR Justification: The AI-generated impacts provide relevant projections for labor equity advancements and international applications; however, these insights remain somewhat generalized and lack the predictive depth evident in the REF submission, particularly in the areas of gender-specific and austerity-driven labor policies.

"
"83. Study Title : Creating Value and Transforming Lives through Arts and Creative Media Practice
Rater 1 Rating = 3
IR Justification: The AI-generated content provides insightful assessments by predicting potential applications and considering broader policy impacts. However, some insights remain surface-level compared to the in-depth contextual examples in the REF submission. For instance, the REF version's discussion on specific changes to planning outcomes and community-led activism offers a deeper understanding of direct impacts on local governance and policy compared to the AI version.
Rater 2 Rating = 4: Very Good
IR Justification: The AI-generated report shows insightful connections, such as how participatory art impacts public perception and policy, extending beyond mere reporting. For instance, it predicts potential benefits for policy advocacy and educational frameworks on heritage conservation and social inclusion. However, the original REF submission's depth in contextual examples lends it slightly greater situational richness.
Rater 3 Rating = 4 (Very Good)
IR Justification: ChatGPT provides some insightful extrapolations on future trends, particularly in scalability and international relevance for other post-industrial or informal urban areas, which are not explicitly covered in the REF version. It suggests wider applicability of community-driven heritage preservation through arts and policy influence. However, some insights specific to local dynamics (e.g., Govan Docks) and named stakeholders in the REF version lend richer detail.
Rater 4 Rating = 4 (Very Good)
IR Justification: The AI-generated impacts demonstrate foresight, particularly in predicting implications for similar global contexts and offering a forward-looking perspective on the potential for the arts to influence policy and social cohesion. It successfully extrapolates on the original themes, suggesting broader societal applications.

"
"84. Study Title : A Transformation in Creep Condition Monitoring for High Temperature, High Pressure Components
Rater 1 Rating = 3 - Good
IR Justification: The AI version offers insightful connections, extending potential impacts to future applications and suggesting scalability across sectors such as renewable energy. However, the REF submission provides additional actionable insights by specifying local applications, policy implications, and contributions to industrial training and regional economic development. The AI's insights lack the same level of actionable specificity but contribute by identifying broader implications beyond the original study.
Rater 2 Rating = 3 (Good)
IR Justification: The AI version provides useful insights, especially regarding technological and commercial impacts, and suggests future applications and scalability of the technology. However, the REF version's insights on real-time challenges, including specific maintenance needs in Eskom and SASOL's use of the technology, contribute deeper contextual relevance that anticipates direct benefits.
Rater 3 Rating = 4 - Very Good
IR Justification: ChatGPT's generated impacts provide insightful implications on the broader effects of WeldCore's technological contributions, including suggestions for future applications in various industries. The content extends some implications to future industrial and environmental impacts, predicting cost savings and enhanced safety, which shows an understanding of the lasting potential of the research. However, it lacks some finer details on SASOL and ESKOM's specific organizational improvements and lacks recognition of regional innovation awards, which added value to the REF submission.
Rater 4 Rating = 5
IR Justification: The AI report introduces potential long-term outcomes of the technology, such as decreased environmental impact and a reduction in energy consumption, which are relevant yet not explicitly discussed in the REF submission. Additionally, the AI-generated report suggests broader economic impacts by outlining future trends and emerging possibilities for

transferability, providing valuable foresight that adds an insightful layer to the existing
impact assessment.

"
"85. Study Title : The creation (and re-creation) of contemporary female heroines at the center
of new plays for the theatre.
Rater 1 Rating = 4 - Very Good
IR Justification: The AI-generated content provides valuable forward-looking insights into the
sustained impact of Harris's work on public perception and future theatre programming practices.
It extends the analysis by suggesting potential long-term effects on gender equity in the arts
globally and the continuity of these impacts within academic and creative fields. However, the
REF version more directly relates current impacts to specific future projects or institutions,
which the AI-generated version does not thoroughly address.
Rater 2 Rating = 3
IR Justification: While the AI-generated impacts offered meaningful insights on the broader
implications of Harris's work for gender representation in the arts, they did not fully capture
the detailed, project-specific context found in the REF submission. For instance, while the REF
submission spoke to specific societal and professional shifts in the Turkish and Kurdish
theatrical contexts, these were less highlighted in the AI analysis. The REF's insights into
changing gender roles across different global contexts demonstrated richer, contextualized
insight.
Rater 3 Rating = 4
IR Justification: The AI content provides insightful analysis by projecting how Harris's works
may inspire future theatrical norms and contribute to cultural discourse on gender
representation. It also extrapolates potential impacts on academic and media discourse regarding
women's roles in the arts, which suggests a forward-looking interpretation of the impact.
Rater 4 Rating = [4 - Very Good]
IR Justification: The AI-generated report introduces relevant anticipations for future trends in
theatre and cultural discourse regarding female representation, suggesting Harris's influence
could extend beyond direct production impact to inspire structural change in theatre programming
internationally. However, it does not capture the mentoring and legacy components fully, such as
her role in mentoring emerging playwrights in Turkey and other countries.

"
"86. Study Title : Changes to cervical screening policies following the rollout of human
papillomavirus (HPV) vaccination
Rater 1 Rating = 3 - Good
IR Justification: ChatGPT's insights extend the REF's findings by noting potential effects on
the long-term economic and public health impact globally, indicating scalability and broader
application. However, it does not predict specific policy advancements or in-depth clinical
impacts that might result from the study, as the REF submission does.
Rater 2 Rating = 4 (Very Good)
IR Justification: The AI report provides insightful extrapolations, suggesting potential policy
adoption pathways and implications for similar vaccination programs globally. Its integration of
predictive analysis offers an advanced understanding of the likely long-term public health
impacts, demonstrating nuanced foresight not fully explored in the original REF submission.
However, the lack of economic projections limits a complete forward-looking view.
Rater 3 Rating = 4 (Very Good)
IR Justification: The AI content shows insightful projections on future policy and clinical
applications, particularly in predicting scalability of HPV vaccine policies and potential
cost-saving benefits in healthcare. However, it sometimes lacks the depth of specific evidence
highlighted in the REF submission.
Rater 4 Rating = 4 (Very Good)
IR Justification: The AI-generated insights extend beyond the original content, especially
regarding potential long-term impacts of improved cervical cancer prevention strategies and the
reduction of disease incidence over time. Additionally, the AI's mention of indirect
beneficiaries (e.g., future populations with reduced HPV prevalence) suggests further-reaching
implications not explicitly stated in the REF submission.

"
"87. Study Title : Combating Crop Losses and Improving Global Food Supplies through Mathematical
Modelling of "Gene Silencing"
Rater 1 Rating = 4 - Very Good
IR Justification: The AI-generated impacts presented an insightful analysis, particularly in
predicting the implications of sustainable crop protection and the global relevance of
biostimulant use. Its content extended into potential future applications and offered a broader

view on biostimulants' role in organic farming. It suggested lasting policy shifts and alternative technological partnerships that would advance sustainable practices, which demonstrates a meaningful extension of ideas beyond the core REF submission.
Rater 2 Rating = 3 (Good)
IR Justification: The AI content provided some predictive insights, especially on scalability and adaptation to different agricultural ecosystems, which add depth. However, it lacked some actionable insights found in the REF submission, particularly around direct applications of RNAi techniques in local pest management practices. REF submission offered more extensive insight into region-specific adoption and anticipated economic shifts in Ukraine's agribusiness.
Rater 3 Rating = 4
IR Justification: The AI-generated impacts are insightful, providing a forward-looking view on scalability and the potential future influence of biostimulants. This includes broader implications for global agriculture beyond the initial study area. However, it missed some depth in stakeholder perspectives compared to the REF submission.
Rater 4 Rating = 5
IR Justification: The AI-generated impacts offer considerable insight by linking the research findings to broader, long-term agricultural sustainability trends and potential policy shifts in pest management. The content projects the research's implications beyond Ukraine, considering the scalability of RNAi-based methods to global agriculture, which could shape future research and implementation.

"
"88. Study Title : Emerging Media, Learning, and Organisational Practice - Driving Change in Tourism and Education in Northern Ireland
Rater 1 Rating = 4 (Very Good)
IR Justification: The AI version demonstrates insight by predicting possible applications in global tourism and education sectors and recognizing the potential of AR as a disruptive innovation in cultural heritage. The AI text anticipates future applications of the research, such as scalability to other cultural sites, providing a broad perspective that aligns with long-term trends. However, the REF version includes specific foresight regarding immediate regional strategies for enhancing tourist experiences and curriculum improvements, reflecting a higher level of local insight.
Rater 2 Rating = 5 - Excellent
IR Justification: The AI-generated impact demonstrates insightful foresight by identifying potential future implications of the research on AR integration and educational technology across various regions, suggesting pathways beyond the current applications. It projects future expansion into global tourism industries, something that extends beyond the REF submission's localized focus, thus adding a layer of visionary analysis. This predictive insight allows a broader understanding of the research's potential evolution, making it an excellent addition to the impact narrative.
Rater 3 Rating = 3 (Good)
IR Justification: The AI-generated impacts include valuable insights on the possible broader implications of AR technologies and learning spaces, suggesting future global applications and interdisciplinary benefits. However, while insightful in scope, it lacks depth in predicting long-term regional educational impacts, as highlighted in the REF.
Rater 4 Rating = 3 (Good)
IR Justification: The AI content offers some insightful interpretations of potential future impacts, particularly in predicting wider applications of AR in tourism and education. However, it lacks the in-depth insights seen in the REF submission regarding the intricacies of stakeholder engagements and strategic decisions (e.g., budgeting, specific user journey design feedback) that influence policy and organizational changes in Northern Ireland's tourism sector. Consequently, while AI-generated impacts are forward-looking, they lack the tangible examples and situational specificity that add depth to REF's impact narratives.

"
"89. Study Title : Global adoption of the Dermatology Life Quality Index (DLQI) into clinical practice
Rater 1 Rating = 4 - Very Good
IR Justification: The ChatGPT impacts show insight by contextualizing DLQI's significance within broader healthcare and patient care strategies, foreseeing applications in health economics and advocating its utility in chronic disease management globally. It provides actionable suggestions for DLQI integration into electronic health records (EHR) and highlights its influence in patient-centered approaches, which is insightful and applicable.
Rater 2 Rating = 4 (Very Good)
IR Justification: The ChatGPT-generated impacts demonstrated insight into potential global applications of DLQI in patient monitoring and economic assessments of dermatological

interventions. Additionally, the AI version suggested future applications, such as expanding DLQI into non-dermatology fields, showing a foresight that adds a layer of insight. However, it lacked detailed foresight on regulatory and licensing structures that were included in the REF submission, which provides more directly actionable insights for policy integration.
Rater 3 Rating = 3 - Good
IR Justification: The AI-generated impacts demonstrate a good level of insight, recognizing the DLQI's potential influence on clinical practices and patient care quality. It also highlights prospective applications in policy and patient-centered healthcare approaches. However, the insights could have extended further into economic sustainability or novel applications within specific health technology frameworks, which would bring it closer to the in-depth foresight shown in the REF submission.
Rater 4 Rating = 4
IR Justification: The AI content demonstrates insight by suggesting future applications, such as the potential of DLQI in broader clinical guidelines and digital health tools. It further connects the research impact with future healthcare modeling and personalized patient care enhancements, aspects only lightly addressed in the REF case. However, the insights lack specific mention of existing or emerging partnerships that could anchor these projections, which are crucial for establishing grounded future directions.

"
"90. Study Title : Improving the healthcare experiences of children and young people
Rater 1 Rating = 4 (Very Good)
IR Justification: The AI-generated content demonstrated insightful interpretations, particularly by predicting future trends in pediatric care, such as increased advocacy for non-invasive interventions and extended impacts on global healthcare practices. These inferences align well with the REF submission's discussions on reducing pediatric distress, although they were presented in a more synthesized manner rather than through step-by-step pathways. The AI content additionally provided forward-looking implications on how the frameworks could influence mental health support protocols for children, reflecting well on potential long-term societal benefits not fully outlined in the REF submission.
Rater 2 Rating = 3 (Good)
IR Justification: The AI content provides valuable insights into potential expansion impacts and suggests broader applicability of interventions, such as further enhancing child health literacy and healthcare provider practices. However, the AI does not fully extend on nuanced insights from the qualitative findings on emotional experiences during medical interactions, which could enhance its depth of insight.
Rater 3 Rating = 4 (Very Good)
IR Justification: The AI-generated impacts present thoughtful connections between the research findings and potential policy shifts, behavioral changes in healthcare professionals, and implications for training. It suggests that these impacts could inform broader healthcare practices, which aligns well with insights from the REF version, though some specific local impacts were less pronounced.
Rater 4 Rating = 3
IR Justification: The AI-generated impacts bring valuable predictions for future application across broader healthcare settings, offering potential for cross-disciplinary influence. However, the REF submission provides a more nuanced analysis of parent-healthcare provider relationships, especially in navigating the complexities of pediatric pain management, an element only partially covered in the AI version. The REF submission also offers concrete suggestions for training and support that deepen its insights on long-term impact.

"
"91. Study Title : The value of the Carers' Alert Thermometer (CAT) in identifying family carers' needs and supporting them in their caring role
Rater 1 Rating = 3
IR Justification: ChatGPT's interpretation adds insightful perspectives on long-term applicability, such as potential scalability, future policy influence, and interdisciplinary relevance. These reflections offer forward-thinking insights, albeit more speculative, while the REF version provides grounded insight based on actual implementations and carer feedback, focusing on existing frameworks and adaptations.
Rater 2 Rating = 5 (Excellent)
IR Justification: The AI-generated report offers a highly insightful view, with forward-looking predictions for CAT's potential role in shaping mental health policy for carers, which adds a layer of depth beyond the immediate impacts in the REF submission. The AI version provides actionable insights into future integration in healthcare systems, scalability, and its role in preemptively reducing carer crises, demonstrating a profound understanding of the research impact's long-term potential.

Rater 3 Rating = 4 (Very Good)
IR Justification: The AI-generated impacts demonstrate considerable insight by addressing potential policy and healthcare system benefits that extend beyond immediate carer needs, such as anticipated trends in carer support integration across health services. ChatGPT's extrapolation of future digital applications for CAT and its connection to broader public health policy suggest a forward-looking approach, adding value to the stated impacts in the REF submission.
Rater 4 Rating = 4 - Very Good
IR Justification: The AI-generated insights predict the future implications of CAT's scalability and adaptability in various caregiving contexts, offering a useful forward-looking perspective, though lacking the same depth of real-world testimonials as in the REF.

"
"92. Study Title : God-complexity and the Multiple God-Aspects Framework
Rater 1 Rating = 3 - Good
IR Justification: The AI-generated content exhibited a degree of insightfulness by extending potential implications of the framework beyond the explicit statements in the REF submission. For example, it speculated on long-term impacts on mental health policies and therapeutic frameworks. However, it missed specific contextual insights found in the REF document, like early-stage clinical applications and unique stakeholder feedback, which could have added depth.
Rater 2 Rating = 4
IR Justification: The AI version introduces insightful potential applications for future studies, offering predictions about its applicability in new contexts, such as broad mental health applications and interdisciplinary impact potential. The insights are largely actionable but lack the specific foresight into potential implications for cultural and societal discourse shifts, as was detailed in the REF submission. Furthermore, the AI version's prediction of potential policy influence aligns well but lacks the depth of prediction tied to specific clinical and cross-cultural applications, as covered in REF.
Rater 3 Rating = 3
IR Justification: The AI-generated impact includes insights on future research implications and potential therapeutic applications. However, it lacks some depth in stakeholder responses and the long-term community impact detailed in the REF submission.
Rater 4 Rating = 4 (Very Good)
IR Justification: The AI-generated impact analysis includes forward-looking predictions regarding potential applications in clinical settings and broader religious communities, providing an insightful perspective that extends beyond what is outlined in the REF report. The AI analysis also speculates on potential applications in public health. This added speculative layer enriches the report but lacks some of the evidence-backed details of the REF submission.

"
"93. Study Title : Abertay Game Lab: play, performance, and public engagement with games
Rater 1 Rating = 4 (Very Good)
IR Justification: The AI-generated version provides insightful interpretations of the impacts by extrapolating the project's long-term influence on cultural, social, and industry sectors. It predicts the evolution of participatory play events and their potential to reshape public engagement with video games. The AI expands upon the data by suggesting applications for educational curricula, inclusive event designs, and possible policy implications—elements that extend beyond the REF submission's specific scope.
Rater 2 Rating = 3
IR Justification: The AI-generated content introduces relevant perspectives on the potential future impacts and scalability of game curation and public engagement through play. This reflects the AI's broader predictions for long-term cultural influence and its emphasis on inclusive practices that may further diversify participant demographics in gaming exhibitions. While insightful, the AI-generated version could have offered deeper predictions based on specific outcomes from the REF submission, such as the sustained impact of public VR installations or the influence on educational methods within game design.
Rater 3 Rating = 3 (Good)
IR Justification: The AI-generated version presented insightful points on the transformative role of interactive media in public spaces. However, the REF document demonstrated higher insight by linking these points directly with observed changes in visitor engagement and institutional collaborations.
Rater 4 Rating = 4
IR Justification: The ChatGPT-generated version provides insightful projections on future industry trends, such as potential impacts of game-making on public interaction in museum and non-traditional gaming spaces. This insight adds value by predicting extended applications of game design principles. Yet, certain insights into regional impact strategies (e.g., VR

applications specific to Scotland's local context) were missed, limiting some localized
insightfulness.

"
"94. Study Title : Improving Care for Patients with Chronic and Distressing Tinnitus through
Mindfulness Based Interventions
Rater 1 Rating = 4 - Very Good
IR Justification: The AI-generated impacts display insightfulness in projecting possible future
applications of MBCT-t, especially for chronic pain and mental health. The AI text discusses
potential long-term cost savings and societal benefits from digital and telemedicine
applications, offering strategic suggestions for enhancing reach. However, while insightful, it
lacks some specific context provided in the REF on changes in patient quality of life and health
economic modeling impacts, which give practical, real-world perspectives.
Rater 2 Rating = 4
IR Justification: The AI content provides valuable insights, discussing future trends like the
broader adoption of MBCT in treating related mental health issues and potential healthcare cost
reductions due to fewer patient re-referrals. This forward-looking perspective extends beyond
the explicit statements in the REF submission. However, further insights could have included
specific predictions regarding MBCT-t integration into global health systems, an element
partially addressed in the REF.
Rater 3 Rating = 5 (Excellent)
IR Justification: The ChatGPT version provides forward-looking insights, such as predicting
future integration in digital health and potential scalability across chronic health conditions,
showing depth beyond the REF submission by speculating on broader applications and long-term
impact.
Rater 4 Rating = 5
IR Justification: The AI-generated report demonstrates high insightfulness by extending
implications of MBCT-t beyond immediate clinical outcomes, suggesting its relevance in mental
health care and in shaping future therapeutic practices for chronic auditory conditions. The
predictions on MBCT-t's role in long-term tinnitus management align with trends in mental health
treatment but also envision MBCT-t's applicability in broader healthcare settings, showcasing
forward-thinking relevance.

"
"95. Study Title : Developing and Embedding Effective Careers Guidance for Young People in
England
Rater 1 Rating = 3 (Good)
IR Justification: The AI version offers a good level of insight, with some additional
connections to long-term educational outcomes like social mobility and workforce development.
However, it lacks some of the unique contextual insights of the REF submission, such as the
strategic involvement of policy institutions and the direct influence on career leader roles,
which add a layer of practical relevance not fully captured in the AI summary.
Rater 2 Rating = 3: Good
IR Justification: The ChatGPT-generated impacts exhibit meaningful insight, especially in
predicting long-term effects on educational outcomes and economic stability by enhancing career
readiness. However, it somewhat generalizes the mechanisms of influence, omitting the specific
steps by which the Gatsby Benchmarks' adoption has transformed practice within UK regions as
detailed in the REF submission. While the AI report projects societal mobility benefits and
anticipates the model's adaptability, it lacks some specificity concerning existing stakeholder
feedback that could substantiate these insights.
Rater 3 Rating = 3 (Good)
IR Justification: The AI-generated version provides actionable insights and suggests the
research's long-term social mobility impacts. However, it lacks predictive insights on the
benchmarks' potential for future policy expansions or deeper integration in non-educational
sectors, which the REF version subtly implies by highlighting strong government support.
Rater 4 Rating = 4
IR Justification: The AI-generated version offered a well-rounded view that provides foresight
into the potential scalability of the career guidance framework, including hypothetical
applications in non-English education systems. It also extrapolated potential policy
advancements based on the benchmarks' adoption trends, which could assist stakeholders in
envisioning long-term impacts. However, the insight did not extend to specific roles and
responsibilities within schools, such as career leaders, as effectively as the REF version.

"
"96. Study Title : Design meets disability: changing the relationship between disability and
design in business, culture, practice, and education

Rater 1 Rating = 4 - Very Good
IR Justification: ChatGPT's analysis provides actionable insights that extend beyond REF
details, such as scalable design principles that may impact broader sectors like healthcare and
education. Predictions about the long-term cultural normalization of disability-inclusive design
suggest deep insight, although some specific strategic details in industry influence are more
explicitly covered in the REF submission.
Rater 2 Rating = 4 (Very Good)
IR Justification: The AI-generated content provides meaningful insights into how inclusive
design can reshape perceptions of disability, potentially impacting various sectors. While not
as specific in industry predictions, it captures future implications of design beyond immediate
applications. The REF submission adds foresight with specific discussions on industry trends and
organizational shifts, like The Valuable 500's commitment to inclusivity and how it inspires a
"trickle-down" effect across industries.
Rater 3 Rating = 5 (Excellent)
IR Justification: The ChatGPT version provided forward-looking insights on how the design
principles could influence sectors like education, policy, and additional industries. The AI
emphasized the innovation potential of treating disability as a unique creative lens, thus
enriching its relevance beyond immediate stakeholders.
Rater 4 Rating = 5
IR Justification: The AI-generated impacts presented actionable insights by suggesting how
inclusive design principles could influence other industries and anticipate future technological
and societal shifts. This was consistent with the REF submission's aim to encourage inclusivity
as a fundamental business and cultural ethos.

"
"97. Study Title : Digital Archiving for Curation and Dissemination
Rater 1 Rating = 4
IR Justification: ChatGPT captures detailed pathways for sustaining impact, such as online
archival access and cross-disciplinary relevance. However, it does not speculate on potential
international policy influence as directly as REF, which discusses shifts in curatorial
standards and long-term policy impact within cultural heritage institutions.
Rater 2 Rating = 4
IR Justification: AI offered insights into cultural accessibility but lacked specific trend
projections or policy impacts noted in REF.
Rater 3 Rating = 5
IR Justification: The AI-generated impacts provide additional insight into the long-term
potential of the digital archive in promoting interdisciplinary research and influencing public
engagement practices. Predictions of future scalability, alongside the potential for new
archival standards, reflect a forward-looking perspective that could influence similar archival
efforts in other regions.
Rater 4 Rating = 4
IR Justification: The AI-generated version demonstrates a high level of insight, particularly in
predicting future impacts and highlighting how preservation techniques can foster long-term
access and sustainability. While REF's case study is highly grounded in specific historical
contexts, the AI-generated version adds future-oriented insights into how such archiving
processes may evolve with technology, enhancing its foresight.

"
"98. Study Title : Prevention and management of head injuries in cricket and rugby union
Rater 1 Rating = 5 - Excellent
IR Justification: The AI impacts provide forward-looking insights on policy implications and
project future directions for injury surveillance adaptations and helmet standards in global
sports contexts. Additional emphasis on non-cricket sports like baseball and potential
innovations in helmet materials (e.g., eco-friendly design) reflect a high level of insight and
an understanding of cross-disciplinary influence. This demonstrates an extrapolated impact that
aligns well with REF goals.
Rater 2 Rating = 3 (Good)
IR Justification: While ChatGPT introduces forward-looking aspects (e.g., potential adoption of
injury surveillance in different sports and settings), the insights are more general and lack
specific actionable steps on emerging injury risk factors. The REF version demonstrates a more
grounded perspective, rooted in actionable and specific measures tailored for cricket and rugby
contexts.
Rater 3 Rating = 5
IR Justification: The AI version demonstrates high insightfulness, particularly by suggesting
future trends in sports safety standards and implications for sports-related health policies.
The AI's projections on technology advancements in helmet design, as well as its consideration

of sustainable materials, contribute actionable insights beyond the REF submission.
Rater 4 Rating = 4
IR Justification: The AI-generated report provided insightful projections about future
applications of the research in other high-contact sports, and the potential for evolving injury
surveillance systems. It presented actionable recommendations based on current findings, like
suggesting updated concussion protocols across multiple sports and sports levels, highlighting
its broader applicability.

"
"99. Study Title : The Kindertransport 1938/1939 to the UK: History Informing the Future
Rater 1 Rating = 4 - Very Good
IR Justification: The AI-generated version provides meaningful insights, such as proposing
future applications of the Kindertransport model to contemporary refugee crises and suggesting
potential impacts on global humanitarian frameworks. These insights extend the implications of
the original REF content, suggesting that historical learnings could inspire improvements in
modern trauma-informed care. However, some forward-looking suggestions lack detailed backing, as
seen in the REF's specific references to past programs that support these suggestions.
Rater 2 Rating = 4 (Very Good)
IR Justification: The AI report provides valuable future-oriented insights, such as potential
applications of historical trauma studies to modern refugee policy frameworks. It offers
additional predictions on the potential for educational impacts and further policy influence,
enriching the discussion. However, the REF submission contains a nuanced cultural and historical
depth that adds specific contextual richness not fully captured in the AI-generated version.
Rater 3 Rating = 3 - Good
IR Justification: ChatGPT's insights are meaningful, particularly in projecting the study's
impact on modern policy debates and educational practices. While the insights align with the REF
version, they occasionally generalize or omit specific future implications discussed in the
original study, such as targeted recommendations for trauma-informed policies.
Rater 4 Rating = 3
IR Justification: While the AI-generated impacts demonstrate some forward-thinking by linking
past events with modern refugee policies, it primarily reflects on existing insights without
offering significant predictions on the policy adaptations needed in response to modern crises.

"
"100. Study Title : An evidence-based approach reduces the local costs of biodiversity
conservation in low- and middle-income countries
Rater 1 Rating = 4 (Very Good)
IR Justification: The AI-generated impacts show a high level of insight, predicting both
immediate and future consequences of biodiversity offset policies and their application in
varied socio-economic contexts. The version delves into how these offsets could reshape
corporate and governmental practices beyond Madagascar, potentially informing international
biodiversity conservation frameworks. However, while it projects future influence on corporate
responsibility, it lacks some immediate stakeholder responses seen in the REF submission.
Rater 2 Rating = 4: Very Good
IR Justification: The AI-generated impacts provided valuable forward-looking insights,
anticipating the scalability and adaptability of conservation strategies to other
biodiversity-rich regions. The generated version also effectively predicted the long-term policy
adjustments required to balance social justice with conservation efforts. However, some nuanced
insights, such as the adaptive challenges faced by policymakers when applying these
recommendations in diverse socio-economic contexts, were less emphasized compared to the REF
version.
Rater 3 Rating = 3 (Good)
IR Justification: The ChatGPT version is moderately insightful, providing actionable
recommendations and discussing potential future impacts on global conservation practices.
However, it does not extend as deeply into future applications or detailed, specific follow-ups
as the REF version, which includes concrete evidence of policy reforms influenced by the
research in real-world contexts.
Rater 4 Rating = 4 (Very Good)
IR Justification: ChatGPT demonstrates insightfulness by suggesting the scalability of Bangor's
research impact on conservation policy frameworks beyond Madagascar, anticipating similar
responses in other regions (e.g., policy adaptations in Latin America or Southeast Asia). The
AI-generated impact also emphasized potential future applications, implying how these principles
could reshape compensation mechanisms for affected communities.

"